

PLUTO AND CHARON: Formation, Seasons, Composition

Michael E. Brown

*Division of Geological and Planetary Sciences, California Institute of Technology,
Pasadena, California 91125; e-mail: mbrown@caltech.edu*

Key Words Kuiper belt, planet formation, outer Solar System

■ **Abstract** Pluto and Charon, once thought to be a singular system in an odd orbit at the edge of the solar system, are now known as members of a vast population of icy bodies beyond Neptune. Models for the occurrence of the odd orbit and formation of these bodies in the context of the total population are reviewed. Pluto's orbital characteristics, coupled with the existence of volatiles on the surface, suggest that large-scale seasonal change should occur on the surface. Models of seasonal variability are discussed, past and current observations are examined for evidence of variability, and a straw-man model of seasonal changes is proposed. Finally, recent observations of the surface composition of Charon are discussed and compared with observations of other similarly sized icy bodies in the outer Solar System.

INTRODUCTION

Pluto, labeled as “not yet explored” on a recent U.S. postage stamp, has been undergoing, in reality, a period of remarkable exploration. Where it was once known only to be a solitary faint yellowish body in an unusual orbit beyond Neptune, it is now known to be a dynamic world with significant surface complexity, large-scale seasonal change, a compositionally distinct satellite, a complicated orbital history, and a vast population of smaller cohorts.

The collection of knowledge about Pluto and its satellite Charon as of 1995 was recently collated into a book containing a series of reviews (Stern & Tholen 1997). This review concerns itself with three areas in which knowledge of Pluto and Charon has expanded significantly in the six years since the writing of that book: the formation of Pluto and Charon, seasonal change on Pluto, and the surface composition of Charon.

Understanding of the formation of Pluto and Charon has progressed recently because of continued observations exploring the Kuiper belt—the vast population of icy bodies outside the orbit of Neptune of which Pluto is a member—and work to understand the formation of these bodies. Section 2 of this review discusses recent models for the occurrence of Pluto's unusual orbits and new simulations of the formation of Pluto and the Kuiper belt from the initial solar nebula.

Progress in determining seasonal evolution of Pluto has been rapid, in part because of new astronomical techniques and instruments brought to bear on the problem, but also, interestingly, from a critical analysis of some of the oldest Pluto data available. Section 3 describes the reasons for expecting large-scale seasonal evolution on Pluto, recent models attempting to describe this evolution, and new observations related to seasonal change. At the end of the section, a new straw-man seasonal model is proposed, with some significant differences from currently accepted views.

Understanding the surface composition of Charon has progressed purely because of observational improvements that now allow direct detailed spectroscopic measurement of the surface. Section 4 discusses four new observations and the view they give of Charon's composition. In addition, new observations of physical properties of Kuiper belt objects (KBOs) allow Charon to be placed in the context of other icy bodies in the outer Solar System.

A final section lists questions remaining about these aspects of Pluto, Charon, and the outer Solar System.

THE FORMATION OF PLUTO, CHARON, AND THEIR COHORTS

Pluto, once thought to be a unique icy planet in a peculiar orbit at the outer edge of the Solar System, is now recognized as the largest (known) body in a vast swarm of KBOs that extend well beyond Neptune. The discovery of Pluto's cohorts has not only made much more clear the story of the formation and evolution of Pluto, but has expanded it into one of the formation and subsequent evolution of the entire Kuiper belt.

The Trans-Neptunian Population

As of June 1, 2001, there were 463 objects known orbiting with semimajor axes beyond Neptune, including Pluto and Charon, in the Kuiper belt [recently reviewed by Jewitt (1999)]. The dynamical properties of the Kuiper belt and their relationship to Pluto are best seen in a plot of semimajor axis versus eccentricity (Figure 1). Pluto fits firmly within the clump of high eccentricity objects with semimajor axes around 39 AU. These objects, collectively known as the plutinos, are all locked into the 3:2 resonance with Neptune. Thus, every 496 years, as Neptune goes around the sun three times, a plutino goes around the sun two times. In addition to the Neptune resonance, Pluto's orbit is inclined out of the plane of the ecliptic by 19 degrees, much higher than any other planet, but well within the realm of the plutinos.

Any of the many plutinos with an eccentricity greater than 0.23 (including Pluto) crosses the orbit of Neptune, potentially leading to close encounters and ejection. Cohen & Hubbard (1965) showed that Pluto is protected from close encounter, as resonant interactions with Neptune force Pluto's perihelion always to occur when Neptune is far from the location of Pluto (see Malhotra & Williams 1997). The other

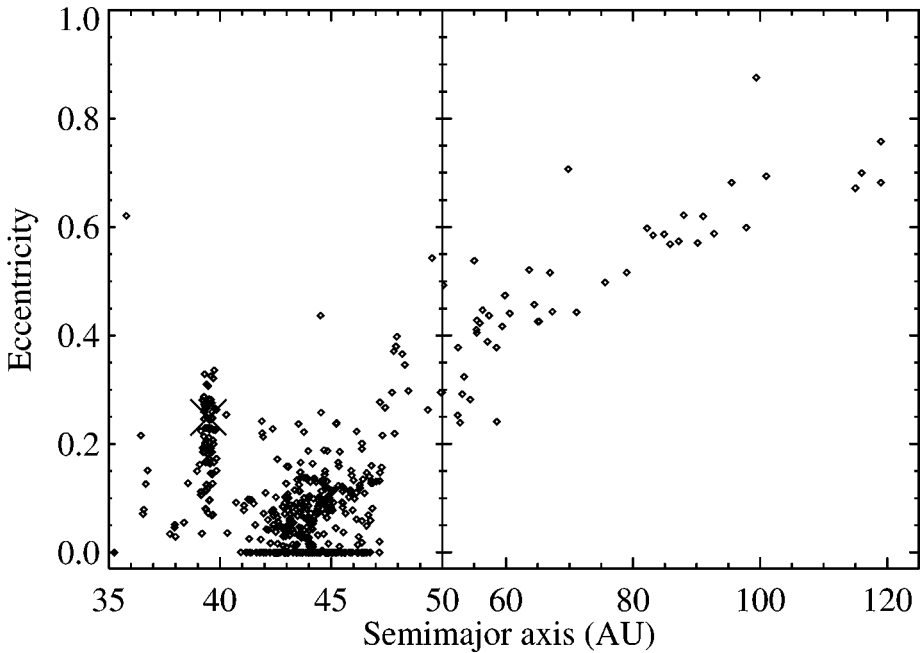


Figure 1 The semimajor axis versus eccentricity of Kuiper belt objects. The location of Pluto is marked with an X. Pluto fits squarely into the class of plutinos that are characterized by semimajor axes around 39 AU and large eccentricities. These objects, like Pluto, are all in the 3:2 resonance with Neptune.

plutinos appear to exist in similarly protected orbits, though the uncertainties in orbital elements of these objects currently make this statement uncertain.

The Origin of the Orbit of Pluto and the Plutinos

Before the existence of the plutinos was known, the orbit of Pluto appeared unique. The fact that its orbit crosses that of Neptune led to the hypothesis, repeated and revised in many forms (see review in Stern et al. 1997a), that Pluto was an escaped satellite of Neptune (Yamamoto 1934, Lyttleton 1936). After it was realized that Pluto was protected from close encounters with Neptune and thus (along with other reasons) could not have been a satellite of Neptune, the only explanation attempted for the peculiar orbit was that Pluto is the last surviving member of a much larger population of like-sized bodies that formed in the outer Solar System. The other bodies were long ago ejected by dynamical instability, but Pluto formed by chance in the dynamically stable 3:2 resonance and thus survived (i.e., Stern 1991, 1992). Now that Pluto is seen as a member of a much larger class of objects, an explanation for their common origin appears necessary.

Malhotra (1993) provided an appealing scenario for the origin of Pluto's orbital characteristics, which extended (Malhotra 1995) to explain the Kuiper belt

and relies on orbital migration of the giant planets and resonant capture of the plutinos. Starting soon after (or perhaps a bit before) the end of giant planet formation, the outer Solar System likely still contained many Earth masses in a remnant disk of planetesimals that had not been swept up into the giant planets. Subsequent encounters with the giant planets scattered these planetesimals either inward or outward and caused the planet to move slightly in the opposite direction. Planetesimals can scatter in either direction, but those scattered outward by Neptune most likely returned to be scattered again in a random direction, whereas some of those scattered inward tended to eventually encounter another planet and thus not return. On average, Neptune scattered more planetesimals inward than outward and its orbit expanded. Saturn and Uranus moved outward similarly, as they received an extra flux of planetesimals from further out. Jupiter, however, is sufficiently massive that planetesimals scattered outward tended to reach distances beyond ~ 3000 AU, where their orbits became modified by galactic tides, and the objects did not return into the inner Solar System. On average, Jupiter scattered planetesimals outward and subsequently moved inward.

As Neptune moved outward, so did the location of its resonances. If Pluto was originally exterior to a resonance, it eventually found itself at a resonance. At that point, if Neptune continued to move slowly enough outward and Pluto's initial eccentricity was sufficiently small, Pluto became trapped in the resonance and moved outward with Neptune. As Neptune moved outward, resonant interactions transferred angular momentum to Pluto, increasing its semimajor axis and causing its eccentricity to increase as $e^2 \approx (1/3) \ln(a_{\text{Neptune}}/a_{\text{Neptune, initial}}) + e_{\text{initial}}^2$. To reach Pluto's 0.25 eccentricity, Neptune needs to have captured it into the 3:2 resonance when Neptune is at ~ 25 AU, suggesting that Pluto's initial semimajor axis was ~ 33 AU.

The resonance capture scenario is extremely attractive because it explains not only Pluto's otherwise apparently unusual orbit, but also predicts that any other small bodies in the outer Solar System would have experienced similar processes. Examination of eccentricity versus semimajor axis of Kuiper belt objects (Figure 1) shows that this prediction is well met. A large collection of objects occupy orbits in the 3:2 resonance along with Pluto. These objects range in eccentricity from approximately 0.08 to 0.34, implying capture by Neptune from 28 to 38 AU, as Neptune moved from 21 to 29 AU. During this movement, Neptune should have also captured objects into other resonances; indeed, the 2:1 resonance has a higher probability of capturing objects. Objects in this resonance are on average 20% farther from the Earth than objects in the 3:2 resonance, so they will be approximately only half as bright, and their detection will be more difficult. For several years no such 2:1 resonance objects were detected, leading to some questioning of the resonance capture theory, but in recent years several objects have been discovered to have orbits consistent with their being in the 2:1 resonance (note that none of the orbits is well enough known to be able to definitively place it inside the resonance).

In addition to the placement of objects in resonant orbits, the resonance capture theory also predicts that there should be no (or few) objects interior to the 3:2

resonance because these objects should have been captured and transported outward. This region of the Kuiper belt is seen as severely depopulated even though these objects would otherwise be the closest and therefore brightest and most easily detected objects in the Kuiper belt. Many potential orbits in this region are stable on timescales of the Solar System (Duncan et al. 1995); no other large dynamically stable region between the orbit of Neptune and the edge of the 2:1 resonance contains so few objects.

The resonance capture theory is compelling for an additional reason: Planetary migration and resonance capture almost certainly must have occurred. Simulations with giant planets and a large number of planetesimals show Neptune migrating by approximately the needed amount (Hahn & Malhotra 1999). Although these computer-intensive simulations did not vary some of the interesting parameters that might affect details of migration, they are in rough agreement with formation models (see below) on the amount of material needed in the early outer solar system. Once such migration is occurring, the objects will be captured unless they either have eccentricities too high or Neptune's migratory motion is jumpy enough that, in a single move, its resonances leapfrog past their resonant objects (note that even if jumpy motion sometimes occurs due to the scattering of an occasional large planetesimal, the Hahn & Malhotra simulations show that the semimajor axis of Neptune is not monotonically increasing—it is more of a generally increasing random walk—thus, any objects stranded by a large leap could easily be recaptured as Neptune wandered inward and then outward again).

Two major aspects of the Kuiper belt are not immediately explained by the resonance capture theory. First, orbital expansion does not affect inclinations, and the inclinations of the plutinos are significantly higher than initially predicted by the theory. In the initial simulations of migration and resonance capture (Malhotra 1995), typical KBO inclinations are only 5° , instead of the 10° and higher inclinations that are typical of plutinos (Brown 2001). Malhotra (1998) and Gomes (2000) show numerically that high inclinations of Pluto and the plutinos can be explained by additional resonances that the objects encounter during migration. For resonantly trapped objects that have had their eccentricities sufficiently excited, orbital expansion can cause capture into the Kozai eccentricity-inclination resonance, which begins to increase the inclination of the object. As Pluto is currently in such a resonance, it is likely that this mechanism operated to increase Pluto's inclination. The ν_{18} resonance, which is the resonance between an object and long-term precessional changes in Neptune's inclination and is currently just beyond the 3:2 resonance with Neptune, would be just inside the 3:2 resonance when Neptune was interior to ~ 34 AU (Gomes 2000). The passing by of this resonance can also lead to inclination increases. A combination of these additional processes appear to be adequate to explain the high inclination excitations of the plutinos in the resonance capture scenario.

The significant eccentricities and inclinations of the nonresonant objects are also not explained by the resonance capture theory. Some increase of eccentricity will occur as weak resonances move past these objects while not capturing them,

but simulations show that this amount of increase is not sufficient to explain the observed distributions (Malhotra 1995). This failure of the resonance capture theory (and the early lack of known objects consistent with the 2:1 resonance) is used as the basis for constructing alternative theories for the formation of the Kuiper belt orbital distributions and, by implication, Pluto's orbit. Most developed is the idea that the Kuiper belt was sculpted by the existence of large scattered planetesimals (LSPs) (Morbidelli & Valsecchi 1997, Petit et al. 1999). The LSP theory suggests that one or more Moon- to Earth-mass planetesimals migrated through the early Kuiper belt before being scattered away by interactions with Neptune. While residing in the Kuiper belt, these LSPs would increase eccentricities and inclinations and cause the removal of a significant fraction of the original belt's members, accounting for the extreme mass depletion of the Kuiper belt (see below). In this scenario, the plutinos are not captured in any special way; they are simply the skeletal remains of the original objects formed in the resonance with Neptune (as in the earlier explanation for Pluto's orbit). This theory, unfortunately, is difficult to evaluate, because the numbers, masses, orbits, and residence times of any LSPs are completely unconstrained. Nonetheless, the LSP theory, too, cannot completely explain, for example, the apparent dual inclination population of the classical KBOs (Brown 2001).

To date, no single theory can explain all dynamical aspects of the Kuiper belt. Rather than suggesting that all theories are wrong, this fact should instead suggest that multiple events must have occurred in the early history of the Solar System. Resonance capture appears to be a natural and almost necessary process in this history. It explains many of the characteristics of the orbits of Pluto, the plutinos, and the overall Kuiper belt, and currently appears the most compelling explanation for the once unique-seeming orbit of Pluto.

The Formation of Pluto and the Kuiper Belt Objects

Before the existence of the Kuiper belt was known, Pluto appeared to be a singular object at the edge of the Solar System, and singular explanations for its formation (i.e., escaped satellite of Neptune, see above) were acceptable. After the discovery that Pluto is protected from close encounters with Neptune (thus could not have escaped) and further, that it has its own satellite, the escape hypotheses were no longer viable; speculation began that Pluto was a member of a now-vanished population that formed beyond Neptune (i.e., Stern 1991). With current knowledge on the existence of a large population in the Kuiper belt, a combined theory of the formation of Pluto and the KBOs now appears both necessary and potentially feasible.

The standard scenario for the formation of planets (or the rock/ice cores of gas giant planets) is that the objects are built by large numbers of random collisions between bodies that then remain gravitationally bound (Lissauer & Stewart 1993). This accretional growth ends when all of the solid material available has been accreted into a few massive objects. Pluto and the Kuiper belt are clearly not in this state, as a large number of small bodies remain unaccreted. Two possible

reasons for such a state could be that accretion is still occurring and insufficient time has passed to allow all of the mass to have accumulated into larger objects, or that, for some reason, accretion was halted at an earlier stage. Stern (1995) ruled out the first possibility; collisions between large KBOs are sufficiently infrequent that, with the current KBO population, 100-km objects (and therefore implicitly also Pluto) cannot be built by pairwise accretion of the current KBO population over the age of the Solar System. Moreover, owing to the large eccentricities and inclinations of KBOs and therefore their high crossing velocities, collisions that do occur tend to be erosive rather than accretional, so collisions make bodies smaller rather than larger. To resolve this dilemma, Stern suggested that the primordial Kuiper belt was both more massive and dynamically colder, so more collisions occurred, and they were gentler and therefore generally accretional.

Following this idea, detailed modeling of accretion in a massive primordial Kuiper belt was performed by Stern (1996), Stern & Colwell (1997a,b) and Kenyon & Luu (1998, 1999). While each model includes different aspects of the relevant physics of accretion, fragmentation, and velocity evolution, the basic results are in approximate agreement. First, with approximately $10 M_E$ (Earth masses) or greater of solid material in an annulus from approximately 35 to 50 AU on very low eccentricity orbits ($e \leq 0.001$), all models naturally produce on the order of a few objects the size of Pluto and approximately the right number of ~ 100 -km KBOs. [Different simulations use different annulus sizes, so direct comparison of the total mass between simulations is difficult. Each will be normalized to a minimum mass solar nebula, defining a minimum mass solar nebula to have a surface density of $\sigma = \Sigma_0(r/r_0)^{-3/2}$, with $r_0 = 1$ AU and $\Sigma_0 = 45 \text{ g cm}^{-3}$. For this minimum mass solar nebula, $25 M_E$ of material are between 35 and 50 AU; thus, the $10 M_E$ of solid material in an annulus is equivalent to $0.4 M_{\text{min}}$.] The models suggest that the majority of mass in the disk is in bodies approximately 10 km and smaller. These bodies suffer frequent collisions and many are ground down to dust over the history of the Solar System. This dust is efficiently removed from the Solar System by Poynting–Robertson drag or photon pressure, thus explaining, at least in part, where the large initial mass has gone. With the number of small bodies diminishing, growth of the larger bodies slows, so the observed KBO size distribution is, in part, a result of this freezing out. In addition, depending on the formation timescale, formation of Neptune begins to induce eccentricities on the KBOs high enough to move them from the accretional to the erosive regime (Stern 1996). This effect also freezes the growth of the largest objects, suggesting that all growth must occur before the ~ 100 Ma growth timescale of Neptune (Lissauer et al. 1995). The initial models by Stern and coworkers produced large objects only on timescales greater than 100 Ma. The most recent model by Kenyon & Luu (1999), which includes the most complete treatment of all of the relevant physics, suggests that Pluto and KBO-sized objects form on 30–40 Ma timescales by standard pairwise accretion as long as the initial mass in an annulus between 32 and 38 AU is between approximately 3 and $30 M_E$ (0.3 to $3.0 M_{\text{min}}$). The large bodies are quickly formed in a runaway growth process after the appearance of ~ 1 -km bodies (Figure 2).

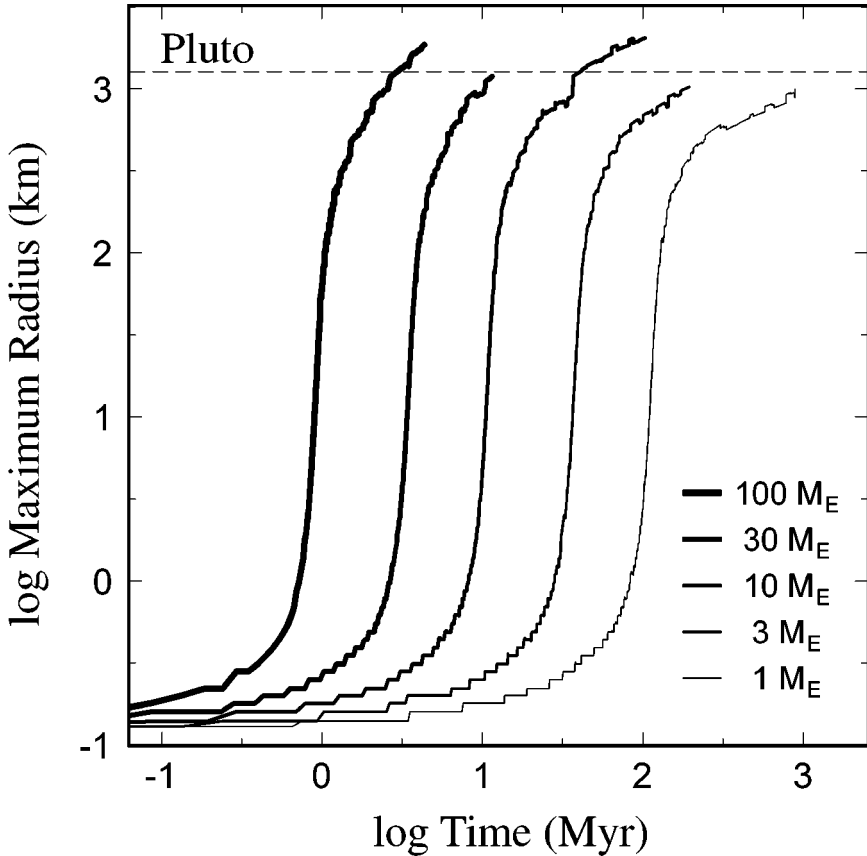


Figure 2 Evolution of the radius of the largest object in the Kuiper belt as a function of initial mass (in Earth masses, M_E . To scale to the fraction of the minimum mass solar nebula, divide by 10.) in an annulus between 32 and 38 AU in the Kenyon & Luu (1999) simulations. The timescale to reach runaway growth and form Pluto-sized objects (marked by a *dashed line*) is a strong function of initial mass. Annuli with masses between ~ 3 and $30 M_E$ can form objects the size of Pluto before the completion of Neptune formation at ~ 100 Ma. Figure from Kenyon & Luu (1999).

Kenyon & Luu suggest that the lack of appearance of this runaway growth phase in previous models is the lack of adequate mass resolution in those simulations.

The current findings suggest that the formation of Pluto (and now also the Kuiper belt) was simply the expected result of extending the ideas about planet formation to the regions immediately beyond Neptune. The fundamental difference between planet formation at Neptune and in the Kuiper belt is simply one of timescale. With the lower masses and longer times between collisions in the Pluto annulus, formation timescales are longer. If left in peace, small objects outside Neptune should

eventually accrete to form objects approaching the size of Neptune's icy core. But the formation of Neptune and its radial migration occurs before such large bodies can develop, and the subsequent increases in eccentricities and inclinations in the Kuiper belt region halts accretion and freezes the size-scale of the largest objects.

Although the general picture of Pluto formation appears solved, a fundamental question remains: How did the initial mass disappear? The most recent accretion models suggest that most of the mass is lost through erosion of small bodies to dust. Dynamical models have long suggested, however, that bodies can be lost due to orbital instability (Duncan et al. 1995). Clearly, both processes operate, but the pictures that emerge from the two mechanisms are fundamentally different. If erosion is the primary loss mechanism, the population of ~ 100 km and above bodies is essentially primordial and has not suffered significant losses. On the other hand, if dynamical losses are the most important, then $\sim 99\%$ of the 100 km and above bodies have been lost, and the Solar System initially contained on the order of 100 Pluto-sized bodies. The accretional models do not generally support such numbers of large bodies without significantly increasing the mass of the solar nebula, though perhaps they have been slightly tuned to create the Solar System that is presently observed. More massive disks are plausible owing to expected inefficiencies in planet formation and the known population of objects such as Oort cloud comets. Indeed, Hahn & Malhotra (1999) suggest that $\sim 50 M_E$ of extra (in addition to the giant planets) material between 10 and 50 AU ($0.6 M_{\min}$) is required to account for the migration of Neptune. Circumstantial (but nonetheless compelling) evidence suggesting the existence of large numbers of large bodies includes the existence of the Pluto-Charon system, the existence of Triton, and the obliquity of Neptune (Stern 1991). The creation of Pluto-Charon likely required the impact of two approximately like-sized bodies that would be the two largest bodies currently known in the Kuiper belt. The probability that the two largest bodies in the belt would collide and create Pluto-Charon is vanishingly small, arguing that many equal-sized bodies must have been present and subsequently vanished. Similarly, the existence of Triton and the large obliquity of Neptune are best explained by the onetime existence of many large bodies being scattered through the Neptune system.

A formation model combining accretion and dynamics addressing these issues is warranted, though computationally extremely difficult. Perhaps the best observational way to address this problem would be a spacecraft observation of cratering rates on Charon or a large Kuiper belt object [Pluto is resurfaced so quickly by frost (see below) that it will not retain the primordial cratering record]. A difference of a factor of 100 in the initial numbers of large objects should be clearly recorded on these surfaces.

SEASONS ON PLUTO

Pluto's high orbital eccentricity, large rotational obliquity, and extreme surface volatility combine to make it the Solar System object most likely to undergo global-scale variation over relatively short time periods. Understanding seasonal

variability of this body requires a coupled understanding of the surface composition and distribution, the atmospheric composition and pressure, and the changes caused by the frequent changes in insolation.

Basic Data

THE ORBIT Pluto's orbit is highly eccentric (possibly caused by resonance capture; see above), and its spin is highly oblique (possibly caused by the Charon-forming impact). These two facts combine to make the surface insolation at any point on the surface highly variable. Averaged over this variation, however, the obliquity of 120° causes greater annual insolation at the poles than at the equator, as each pole spends long periods of time in continuous sunlight. In Pluto's current orientation, which varies on timescales of 10^6 years (Malhotra & Williams 1997), the Sun crosses the equator at perihelion and aphelion, meaning that while the equator receives less average insolation, it also receives the highest instantaneous insolation (Figure 3). This complicated surface insolation pattern has the potential to drive complicated patterns of volatile evaporation and deposition, and potentially leads to global-scale resurfacing on timescales as short as one Pluto year (248 Earth years). Precisely what happens depends on the details of the materials that are presently on the surface of the planet.

THE SURFACE COMPOSITION All of the detailed knowledge of the surface composition of Pluto has been obtained from telescopic observations of the spectrum of sunlight reflected from Pluto's surface. Particularly in the infrared, these reflectance

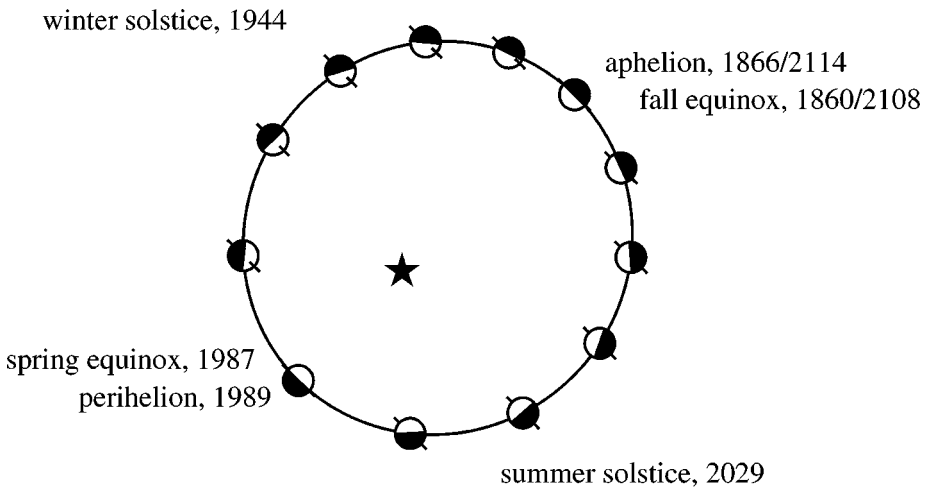


Figure 3 Seasons on Pluto. The position of Pluto is shown every 20.67 years (a Pluto “month”). The seasons are defined for the northern hemisphere of Pluto. Note that Pluto’s “northern” hemisphere is in the south. Pluto observers use the right-hand rule to define north on Pluto, rather than the IAU convention.

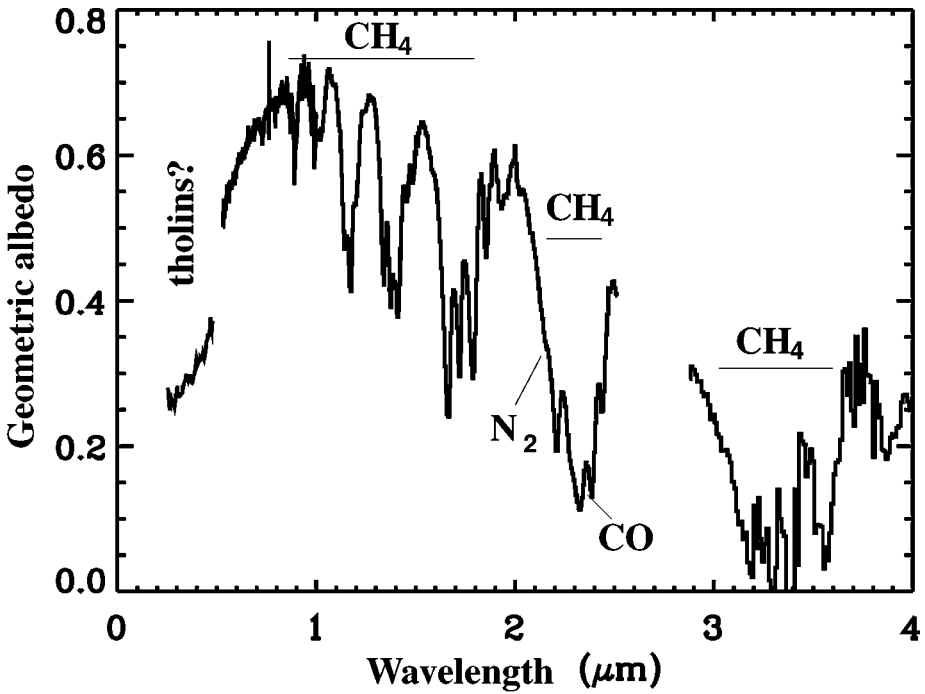


Figure 4 The reflectance spectrum of Pluto from 0.2 to 4.0 μm . The region from 0.2 to 0.5 μm is from Trafton & Stern (1996), 0.5 to 1.0 μm is from Grundy & Fink (1996), 1.0 to 2.5 μm is from Brown & Calvin (2000), and the spectrum to 4.0 μm is unpublished data of M. E. Brown. The absolute reflectances are scaled from Dumas et al. (2001) and Spencer et al. (1990). The major absorption features caused by CH_4 , CO , and N_2 ices are labeled, though at the moderate spectral resolution of these spectra the CO and N_2 features are difficult to discern.

spectra contain absorptions due to the primary surface ices (Brown & Cruikshank 1997). A full review of surface spectroscopy of Pluto is given in Cruikshank et al. (1997). Figure 4 contains a composite reflectance spectrum from the ultraviolet (0.2 μm) to the near infrared (4 μm) showing all of the major absorption features in the spectrum of Pluto. The spectrum throughout is dominated by absorption due to methane, which was the first ice discovered on the surface (Cruikshank et al. 1976). More recently, the absorptions due to N_2 and CO were resolved (Owen et al. 1993).

The detection of these ices immediately suggests several implications. The first is that, because the N_2 molecule has no dipole moment, and thus the 2.148- μm absorption strength is low, the appearance of this band at all suggests an extremely high abundance of N_2 , perhaps dominating the surface. The second implication arises because all of the detected ices are volatile at the temperature of Pluto, with the volatility of N_2 greater than CO , which is greater than CH_4 . Because of this surface volatility, Pluto will be surrounded by an atmosphere; because of the high

abundance and volatility of N_2 , this atmosphere should be dominated by N_2 . The existence of an atmosphere with a mean molecular weight consistent with N_2 was confirmed by observations of a stellar occultation in 1988 (Hubbard et al. 1988, Elliot et al. 1989). The volatility of the surface, combined with the extreme changes in insolation described above, suggests that large-scale surface changes are likely to occur. The final implication of the ices on the surface comes from the fact that no involatile ices have been definitively detected. Water ice, which is essentially like a rock at these temperatures, must be abundant in the interior of Pluto, from both the measured density and from cosmochemical considerations. Water ice and CO_2 , another abundant nonvolatile ice, have been detected on Triton (Cruikshank et al. 1993). Yet neither has been definitively detected on Pluto (detection of water ice is difficult owing to the presence of the CH_4 lines, and its existence has been periodically reported but never published). The volatiles on the surface of Pluto must completely (or at least mostly) cover the rocky interior materials that show through on Triton.

The other obvious spectral feature on Pluto is the red slope in the visible region of the spectrum. All of the ices described above are essentially transparent in the visible, so an additional absorber is required. A persistent hypothesis for the reddening agent on the surface is the existence of tholins, complex organic molecules that can be formed from photolysis of simple organic materials. Unfortunately, these materials have no specific absorptions in these spectral regions, only a generally red color (Khare et al. 1984). Creation of tholins has been performed under experimental conditions appropriate for gas-phase chemistry at Titan and for ice-phase chemistry on comets, but no experiments have been performed using the CH_4/N_2 ice chemistry appropriate for Pluto. It is assumed that the general trends seen in the other experiments will persist, but experimental study of these reactions is critical for understanding surface composition and chemistry on Pluto.

Increasingly detailed spectroscopy and laboratory measurement has led to increasingly detailed knowledge about the state and distribution of materials on the surface. The first hint of the complex state on the surface was the realization that the methane absorption features are shifted by approximately $0.02 \mu m$ from their laboratory positions. Such a shift is seen in laboratory spectra of methane when small amounts are dissolved in N_2 ice (Quirico et al. 1996). With methane dissolved in nitrogen, the methane abundance would be extremely small, on the order of a few percent (Owen et al. 1993). Rotationally resolved spectra in the visible (Grundy & Fink 1996; see below) and higher resolution spectra in the infrared (Douté et al. 1999), however, show that both methane diluted in nitrogen and relatively pure methane exist on the surface. Detailed modeling suggests that their surface coverage abundances are approximately equal, implying that the surface methane abundance is much higher than previously thought.

Little is known about the state of CO. Even if diluted in nitrogen ice, the CO absorption shifts by only $0.0006 \mu m$ (Quirico & Schmitt 1997), which is not resolvable in any of the present Pluto spectra. CO is, however, almost completely miscible in nitrogen, so it should be completely dissolved.

First-Order Seasonal Variability

GENERAL PRINCIPLES From knowledge of the orbit and the surface materials, a first-order understanding of the seasonal variability on Pluto is discerned [see Spencer et al. (1997) and Trafton et al. (1997)]. Important principles that should control the overall seasonal variability are discussed in brief here.

Nitrogen on Pluto is volatile and abundant, leading to buffering of all exposed nitrogen patches at a common temperature. Any patch of nitrogen that is colder will attract more atmospheric nitrogen to condense on it, which will warm it from the excess latent heat. Conversely, any nitrogen patch that is warmer will evaporate more nitrogen until it is cooled by the excess latent heat. As long as atmospheric transportation is faster than the timescale of insolation changes, which should generally hold at the slow rotations of Pluto, the nitrogen temperature will remain constant.

In light of this buffering, any region receiving an excess of insolation will lose nitrogen ice to regions with a deficit of insolation. At perihelion, when the subsolar point is essentially at the equator, this state implies net transport of nitrogen ice from the equator to the poles. On average, however, the equator receives less insolation than the poles, so over the whole orbit, nitrogen will be removed from the poles and deposited at the equator.

Methane, which is less volatile than nitrogen or CO, will move less readily, possibly leading to long-lived lag deposits. Methane deposits have the potential to darken and redden because of the action of UV and cosmic-ray radiolysis and the creation of tholins (Strazzulla & Palumbo 1998). The precise manner in which N₂ sublimates and leaves behind methane is uncertain (Spencer et al. 1997, Trafton et al. 1997), but this process appears to readily explain the existence of both pure and diluted patches of methane on the surface.

MODELS OF SEASONAL VARIABILITY Taking the above principles into consideration, a guess can be made of what Pluto should have looked like in the 1990s when images first became available from the mutual events and HST imaging (see below). Northern winter solstice had occurred in 1944 and spring equinox was in 1987 (Figure 3). The 1990s were a time of rapid movement of the subsolar point from the south to the north, but where the southern pole had been in continuous sunlight for over the past 100 years. Thus, all south polar nitrogen is expected to have been depleted, perhaps with a reddening lag of methane left behind. The nitrogen should have been transported to the north polar regions, causing large bright polar caps that would just be beginning to come into view of the Earth (and beginning to sublime) with the passage of equinox. The equator was experiencing its highest rates of insolation, driving volatiles north and south, but because the equator receives, on average, lower insolation than the poles, a large reservoir of volatiles could still remain.

Detailed modeling of seasonal temperatures, atmospheric pressures, and frost transport was carried out by Hansen & Paige (1996), who considered only an atmosphere and frost covering of N₂ over an involatile substrate. The free parameters

of the model are the thermal inertia and albedo of the substrate, the emissivity and albedo of the frost, and the total N_2 inventory. The thermal inertia of the substrate is a key controlling parameter in the seasonal behavior. For the highest values of thermal inertia, the substrate temperature is controlled by the long-term averaged insolation; thus, frosts build up and stay at the equator and do not experience significant transportation to the poles. For the lowest values of thermal inertia, polar frost deposits come and go quickly, tracking the subsolar point.

The thermal inertia has been estimated from the ISO observations of the thermal light curve on Pluto (Lellouch et al. 2000), which suggest a value of $(1.5 - 10) \times 10^4$ erg cm⁻² s^{-1/2} K⁻¹ [which corresponds to $(0.3 - 2.4) \times 10^{-3}$ cal cm⁻² s^{-1/2} K⁻¹ in the units used by Hansen & Paige (1996)]. This value is in the range of the lowest values considered by Hansen & Paige (1996) and immediately suggests that seasonal cycles on Pluto should respond quickly to changes in insolation. Lellouch et al. point out, however, that the thermal inertia measured from the light curve depends on the properties of the surface only to the depth of the diurnal thermal wave, which will be on the order of 20 cm, whereas the thermal inertia important for seasons depends on the properties of the surface to the much larger depth of the annual thermal wave, which has a depth on the order of 10 m. If a shallow porous layer lies above a more compact layer, the seasonal value of thermal inertia would be larger than that measured from the thermal light curve and could be as high as moderate values of about 3×10^5 erg cm⁻² s^{-1/2} K⁻¹, preferred by Hansen & Paige (1996).

For these moderate values of the thermal inertia, polar caps come and go with the seasons, with a lag depending on the magnitude of the inertia. In all cases, by perihelion passage and equinox in the 1990s, the southern polar cap has completely sublimed and the frost has all moved to the recently dark north (Figure 5). In cases with large enough volatile abundances, however (which might be most relevant, given no spectroscopic evidence of involatile substrate), permanent equatorial frost bands can form in addition to the seasonal caps.

No models have considered the additional effect of CH_4 as a volatile or darkening agent. Until such models are developed, the ideas developed above can be used to speculate on the effects. First, the rapidly receding N_2 polar caps should leave behind a seasonal lag deposit of CH_4 that will potentially darken and redden with time. The equatorial permanently frost-free zones could be the location of tholin and pure methane deposits, or alternatively, if permanent equatorial frost bands exist, these would frequently switch between mass losing, methane-dominated surfaces and mass gaining, nitrogen-dominated surfaces.

Observations of Seasonal Variability

Various observations could potentially distinguish between different seasonal variability models. Different models predict different global albedos as a function of time, with the moderate thermal inertia cases predicting sharp increases in global albedo as polar caps that formed in seasonally dark regions rotate into view and

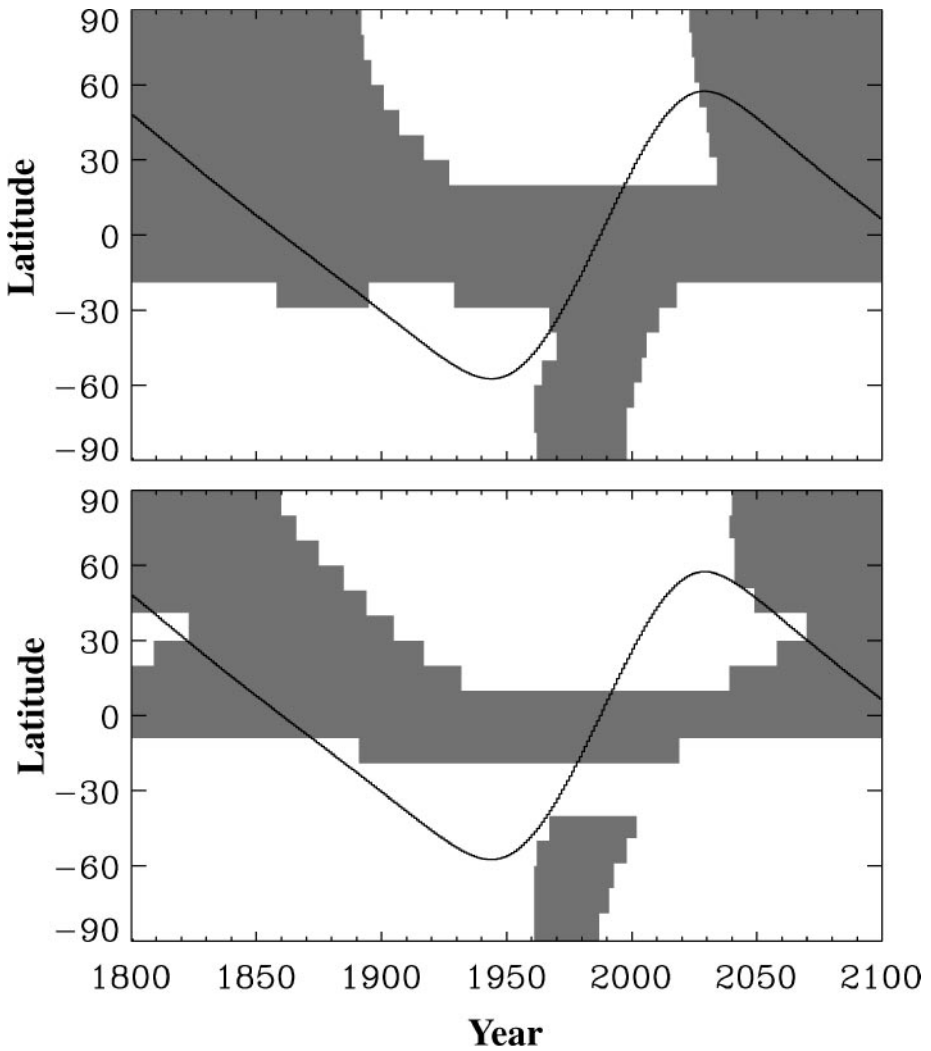


Figure 5 Latitudinal frost distributions as a function of date from two runs of the Hansen & Paige (1996) seasonal frost models. The white regions are covered by pure nitrogen frost, whereas the dark regions are an involatile substrate. The solid line indicates the subsolar point. The top model shows the Hansen & Paige preferred case (run 12) with moderate thermal inertia and N_2 abundance. The bottom model shows a case (38) with increased N_2 abundance where permanent equatorial frost bands form.

only slowly sublime. In addition, the models predict large variations in atmospheric pressure as caps come and go and global temperatures rise and fall. Finally, the models predict different global distributions of volatile ices with season.

PHOTOMETRY Observations definitively establishing changes in the Pluto system with changes in the heliocentric distance or pole position are difficult. Expected timescales of change are decades, whereas many of the observing techniques used to study Pluto are significantly less old, so no suitable baseline exists. One exception to this lack of long-term data is photometry. In the early 1950s, astronomers began undertaking photoelectric photometry of Pluto (see review by Marcialis 1997), and in 1954 Walker & Hardie (1955) obtained the first photoelectric light curve clearly showing Pluto's 6.4-day rotation period. Since that time, new light curves have been obtained on an essentially decadal timescale (see Marcialis 1988), with the post-1980 photometry deriving from CCD rather than photoelectric data. Figure 6 shows the evolution of the light curve from the 1950s to the present. Two important effects are clear: The total albedo of the planet has decreased since 1954 by an average of approximately 25% (with a possible slight increase between 1981 and

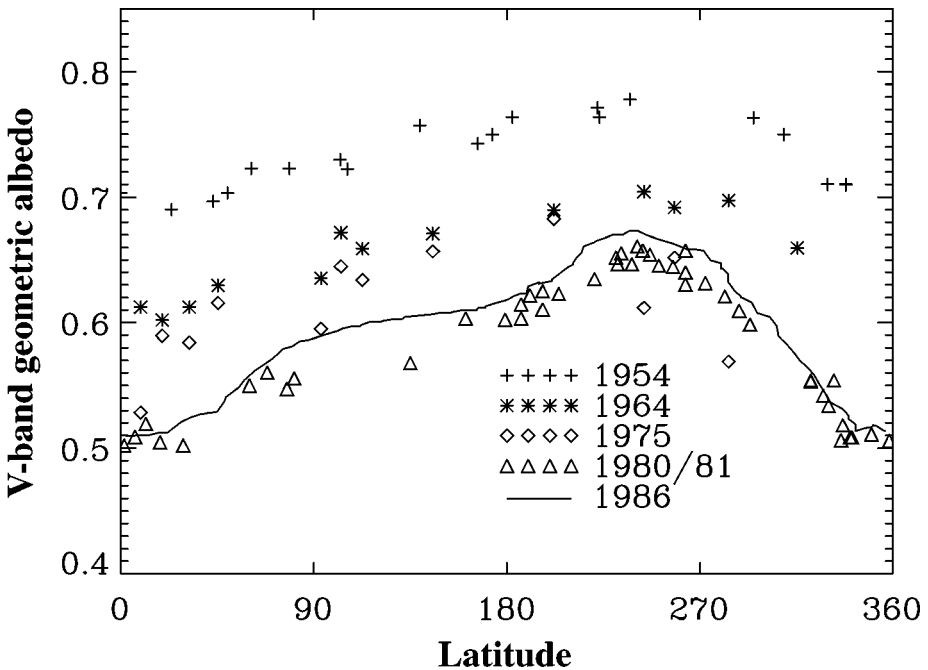


Figure 6 The geometric albedo of Pluto as a function of subearth latitude from 1954 until 1986, from data compiled by Drish et al. (1995). Since 1954, Pluto has gotten darker, but its light curve modulation has increased.

1986), and the magnitude of the light curve modulation has increased in that time period from 13% to 32%. The obvious interpretation of the dimming and light curve modulation increase of Pluto would be that in 1954, when Pluto was being viewed close to south pole-on, a bright polar cap covered most of the visible planet. As time passed, and more northerly latitudes came into view, darker and more variable regions appeared. Unfortunately, this obvious interpretation cannot be correct. Drish et al. (1995) show that no static pattern of albedo on the surface of Pluto is capable of explaining all of the light curve data. Although their analysis is primarily devoted to a proper matrix inversion of the light curve data and shows that no mathematically good fit to the entire data set can be obtained, they also show very simply that the light curve data are inconsistent with a static surface. The essential discrepancy is that the magnitude of the dimming from the 1950s to the early 1980s coupled with the relatively high albedo at all times is only possible if a large south polar cap of essentially unit albedo is disappearing from view as the northern latitudes, which must therefore have low albedos, are coming into view. Yet these same northern latitudes that are coming into view must be capable of modulating the total light from the planet by 30%, and this large modulation must not have been visible in 1954. Furthermore, the slight brightening in 1986 occurs when only a small portion of the northern hemisphere is being uncovered; thus, an extremely bright northern cap must also exist. The geometric problem is overconstrained, and no combination of static albedos on the surface of the planet can adequately explain all of these effects.

An attractive hypothesis is that the dimming since 1955 is due not only to the south polar cap hiding from view, but also to the south polar cap subliming away. In this case, both the northern latitudes rotating into view (which have been in darkness for more than a hundred years) and the equatorial regions can be bright. The Hansen & Paige (1996) seasonal variability model with moderate thermal inertia and moderate nitrogen abundance predicts such behavior (Figure 5). The moderate thermal inertia ensures that the dissipation of the southern polar cap began only after southern summer solstice and was not completed until 1970. Meanwhile, a bright northern cap had already formed well before 1950 and would not begin to dissipate until 2020. Unfortunately, this geometric problem is underconstrained. Once albedo features of the surface are allowed to change, no unique reconstruction of the past history is possible. Nonetheless, the analysis of Drish et al. (1995) effectively rules out any static history and demonstrates that large-scale albedo changes must indeed have occurred on the surface of Pluto.

MUTUAL EVENT MAPPING While photometry cannot uniquely constrain seasonal history, it can be combined with instantaneous snapshots of the albedo distribution to help reconstruct history. The first detailed views of the surface of Pluto came from reconstructions of the photometry of the series of Pluto-Charon occultations, eclipses, and transits (“mutual eclipses”) that occurred between 1985 and 1990. Surface mapping was possible because as Charon moved in front of Pluto and its shadow and/or disk blocked Pluto’s surface, progressively different regions of

Pluto were blocked out. Careful photometry and mathematical reconstruction of the light curves from these events allows a limited type of mapping of the sub-Charon hemisphere of Pluto. A complete review and comparison of the different methods and results appear in Buie et al. (1997). A recent reanalysis (Young et al. 1999) combines the largest number of mutual events to obtain the highest possible resolution on the surface. The final map suggests several large-scale features. First is a large dark longitudinal stripe centered at $\sim S12^\circ$ with albedos below approximately 40%. The second large-scale feature is a large region (or perhaps regions) of albedo higher than 75% south of $S30^\circ$. Following earlier interpretations of mutual event data, Young et al. characterize this bright spot as a “bright southern cap,” though they demonstrate that their ability to map features in the far south is extremely limited. The actual southern extent of this feature is unknown, as is the longitudinal range. No similarly bright northern region is seen, again with the same caveats on mapping the extreme northern hemisphere.

Buie et al. (1992) combined mutual event maps with rotational light curves in an attempt to map the entire surface, using a constraint of maximum entropy in the reconstruction. This method requires the selection of a default image, deviations from which define the (negative of the) entropy. Image characteristics unconstrained by data, such as the latitudinal distribution of albedos on the anti-Charon side, depend completely on the default images selected. Buie et al. chose an iterative default image that minimizes local variations in the reconstructed map; thus, the resulting map of the latitudinal distribution of the anti-Charon side will tend to blend smoothly with the well-constrained latitudinal distribution on the sub-Charon side. This choice of default images essentially makes the (not unreasonable) assumption that the unseen anti-Charon side resembles the sub-Charon side in latitudinal albedo distribution. Unsurprisingly, the final map looks roughly symmetric with latitude and thus appears to reinforce the conclusion of the bright southern polar cap seen in the mutual event sub-Charon map.

If taken literally, these maps appear to imply that, as late as 1990, a bright southern polar cap was still in existence and that no northern cap had yet appeared. Such a configuration is difficult to explain; by 1990 the southern hemisphere had been in continuous sunlight for over 100 years, and keeping a polar frost cap in light of this extreme insolation is hard.

DIRECT IMAGING The mutual events provided the opportunity for a snapshot of a fraction of Pluto’s surface, but the opportunity will not be repeated for a long time. Mapping the full disk of Pluto, and mapping it at arbitrary times, requires a different technique. Unfortunately, the disk of Pluto is small. Even at perihelion, it subtends only approximately 0.1 arcseconds. Stern et al. (1997a) obtained images of Pluto in 1994 using the Faint Object Camera at ultraviolet wavelengths on the Hubble Space Telescope (HST), thus obtaining the highest spatial resolution possible on that telescope. From three to four resolution elements are seen across Pluto’s disk in the images. Four images were obtained at approximately 90° separation, therefore covering (at least sparsely) the entire disk. In the subsequent analysis, Stern et al.

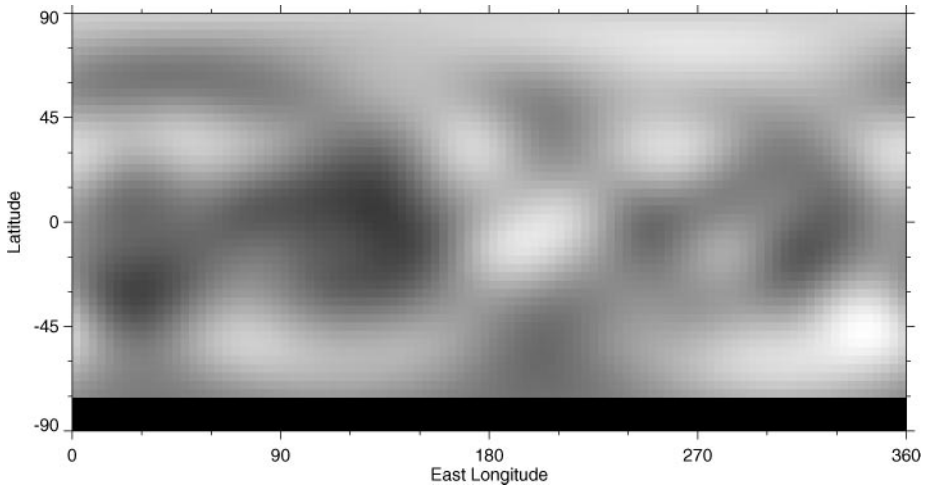


Figure 7 An HST albedo map of Pluto at $0.41 \mu\text{m}$ from Stern et al. (1997a). The map shows that the darkest regions of the surface are roughly confined to the equatorial areas, whereas the brightest regions are widely distributed. The very brightest regions of the planet are at the equator, while no evidence exists for a uniformly bright southern polar cap. The significance level of individual features can be discerned from the original figure in Stern et al.

carefully examined the individual images to verify the reality of features seen, and then constructed a best-fit map of the entire surface by combining the images (Figure 7).

In comparing the HST map of the sub-Charon hemisphere to the mutual event maps, many of the details are reproduced, thus increasing confidence in both methods. Several discrepancies exist, however, predominantly in the polar regions, which are less well mapped by both methods. The HST map shows a more consistently bright northern hemisphere than southern hemisphere, whereas the mutual event maps show the opposite. Comparisons can also be made with the attempt by Buie et al. (1992) at whole-disk mapping; the HST map tells an entirely different story from the generally latitudinally symmetric Buie et al. map. The brightest parts of the anti-Charon side are at the equator rather than the poles; the southern polar region contains large dark areas. Thus, the Buie et al. selection of default images favoring symmetry with latitude gives spurious results on the unconstrained side of Pluto. Whereas Stern et al. (1997a) suggest caution as to which of the mapping efforts to believe, it appears most reasonable, barring specific reasons to the contrary, to trust the most direct method, which would be the HST imaging.

The HST images suggest that the mutual event mapping, while correct, was misleading. The south polar cap apparent in the mutual event maps is seen to be a longitudinally confined spot rather than a band. The dark equator is also longitudinally limited. The brightest spot on the entire planet is on the equator on the

anti-Charon side. No evidence exists for a globally dark equator or globally bright poles. Indeed, the lack of obvious latitudinal correlations with albedo makes the above discussion of expected seasonal effects on albedo seem simplistic. Nonetheless, Stern et al. show that most of the low albedo regions on the planet are confined to the equator, whereas bright regions can be found anywhere. Clearly, albedo on Pluto is much more than a simple seasonal effect that can be straightforwardly inverted to understand seasonal changes. However, the HST map requires that seasonal changes have occurred; with the bright northern hemisphere, it comes nowhere close to explaining the secular dimming of Pluto from the 1950s to the 1980s.

ROTATIONALLY RESOLVED SPECTROSCOPY One method to attempt to understand the causes of albedo variation on Pluto is to compare spectra obtained at different rotational phases of the planet to see if the spectra change with changes in brightness. Buie & Fink (1987) first showed that such an effect exists in the visible region of the spectrum. Grundy & Fink (1996) expanded on the early efforts and compared 20 spectra in the 0.5- to 1.0- μm region of Pluto obtained from 1980 to 1994 and looked at the variation of the methane absorption lines with rotation. They find that the CH_4 band depths peak approximately midway between the minimum and maximum points of the light curve (Pluto longitude of $\sim 340^\circ$). Grundy & Fink develop a surface composition model based on the Buie et al. (1992) maximum entropy surface reconstruction, which we now know to be faulty on the anti-Charon side. Later work uses the HST map instead and is discussed below.

Rotational modulation of the infrared methane bands has also been observed (Marcialis & Lebofsky 1991). Grundy & Buie (2001) obtained a densely sampled collection of infrared spectra from 83 nights between 1995 and 1998. A rough correlation exists between the absorption strengths of the strong CH_4 , CO, and N_2 bands; the amount of dilution of CH_4 ; and the V-band brightness (Figure 8). The absorption strength of the weak CH_4 bands is roughly anticorrelated, as noted in the visible spectra above. Grundy & Buie (2001) argue that the weak CH_4 absorptions are actually a better tracer of CH_4 concentration; large path lengths of methane are required for these absorptions to be visible, whereas the strong bands will show up even when small amounts of methane are diluted in nitrogen. The anticorrelation between the weak methane absorption and the methane dilution strongly supports this point of view.

A first-order interpretation of the rotational spectra is that the bright regions of Pluto are controlled by the presence of N_2 frost, based on the strong correlation between the N_2 absorption strength and the visible light curve. The good correlation of the strong CH_4 bands and CH_4 dilution with N_2 suggests, as mentioned above, that these come primarily from CH_4 diluted in N_2 , so they closely track the N_2 . A similar argument can be made for CO; it is likely primarily found diluted in N_2 (as argued above). Interestingly, the CO rotational modulation appears to differ slightly from that of N_2 , leading Grundy & Buie to suggest that perhaps patches of higher CO to N_2 concentration exist. Whereas Figure 8 is suggestive of this effect,

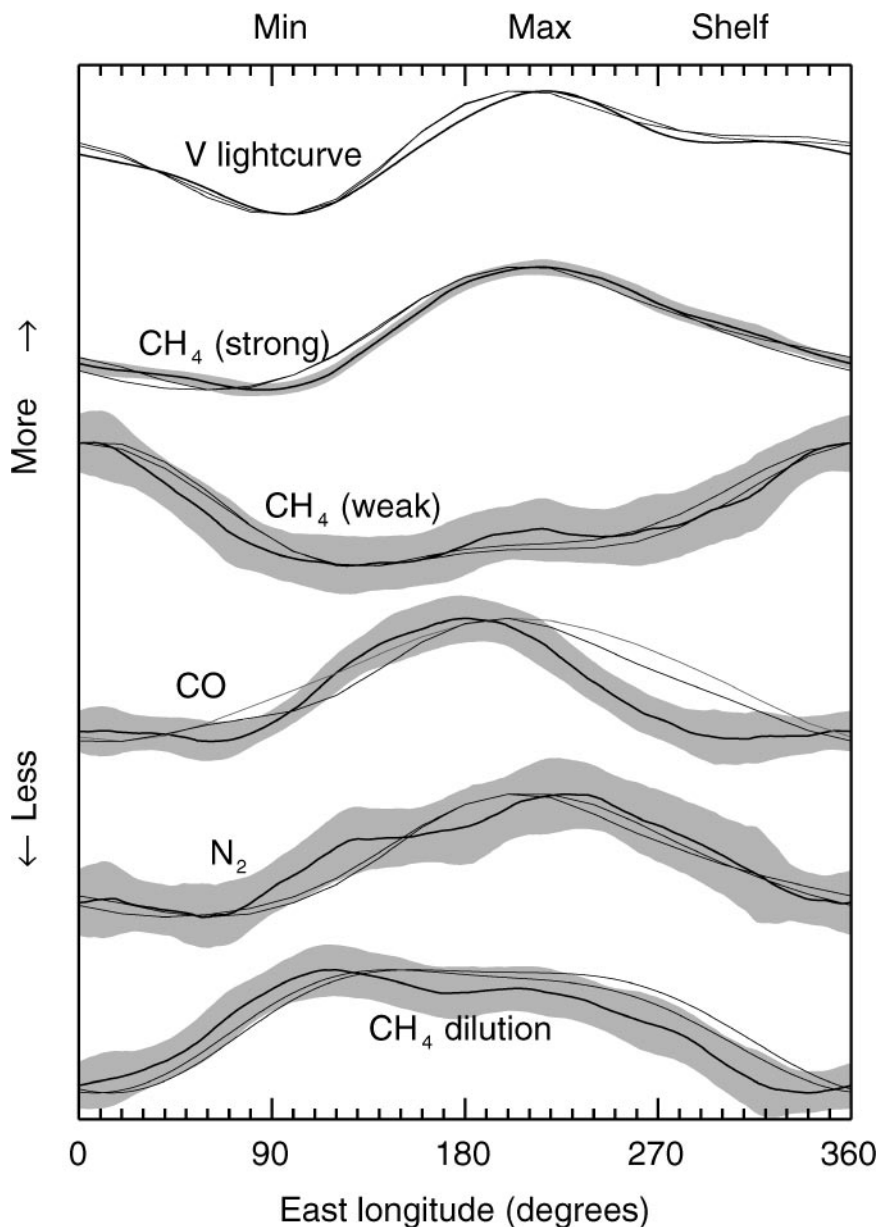


Figure 8 Rotational modulations seen in the spectrum of Pluto, from Grundy & Buie (2001). The strong CH₄, CO, and N₂ absorptions and the amount of CH₄ dilution in N₂ all roughly correlate with the visible light curve. The weak CH₄ absorptions roughly anticorrelate. These facts suggest that at least two separate compositional units exist on the surface of Pluto: a bright N₂-rich unit that contains diluted CH₄ and CO, and a darker CH₄-rich unit. The dashed lines show the predicted rotational modulation from the modified Grundy & Buie compositional models using albedo as a proxy for composition.

examination of the noisy raw data on which this figure was constructed [Grundy & Buie (2001), Figure 6] makes the case less convincing.

Grundy & Buie (2001) developed a model of the surface composition distribution that can fit the observed photometric and spectroscopic modulations. They base their model on a new (currently unpublished) reduction of the HST maps and divide the terrain into three types based on albedo. The lowest albedo regions are composed of a tholin-water ice mix with a spectrum similar to that of the centaur Pholus (Cruikshank et al. 1998) and can be thought of as lag deposits left after complete evaporation of the N_2 and CH_4 ices. This unit appears predominantly in the southern hemisphere and equatorial regions, which is where insolation has been highest for the past 100 years. The mid-albedo regions, typically proximate to the low albedo regions, are assumed to be CH_4 with minor contamination of tholins, and could be regions where sublimation of N_2 has left behind the less volatile CH_4 ices. The highest albedo regions are N_2 ice with diluted CH_4 and CO and represent areas of recent frost deposition. This model (Figure 9) provides a moderately good fit to the spectral rotational modulation seen on Pluto. Grundy & Buie (2001) modify the model slightly by hand to make the fit better by allowing some of the areas to be a checkerboard mix of two units that still give the same albedo (Figure 9). An even better fit can be had in this case, as seen in Figure 8. One obvious problem with this relatively simple model is that it does not distinguish between regions that, by eye, appear quite distinct. For example, the brightest regions of the planet are located in several features along the equator, yet these are all lumped together with the polar regions and labeled as nitrogen rich. It would be interesting to explore the implication that the polar regions—or at least the south polar regions—are not nitrogen rich. The misleading south polar bright spot seen in the mutual event maps continues to lead modelers to expect a bright (and therefore nitrogen-rich) south polar cap when the evidence for one is perhaps no longer compelling. An additional difficulty is the assumption that there is a one-to-one correspondence between albedo and compositional units. Although such an assumption is a reasonable starting point, recent evidence from multicolor mutual event maps (Young et al. 2001) suggests that the dark equatorial region on the sub-Charon side has large color—and therefore presumably also composition—variations across it. Nonetheless, the Grundy & Buie models provide the best starting point for constructing a composition-seasonal variability model of the planet.

A New Straw-Man Seasonal Variability Model

Using the observations and modeling above, a new straw-man view of Pluto's seasonal variability is constructed. The problem is severely underconstrained, so new observations will continue to update or even significantly change this picture.

The first constraint used here is the Pluto photometry from 1955 on. As discussed above, these data can be explained if a bright southern cap existed in 1955 but has mostly sublimed away since then, and the northern and equatorial regions that

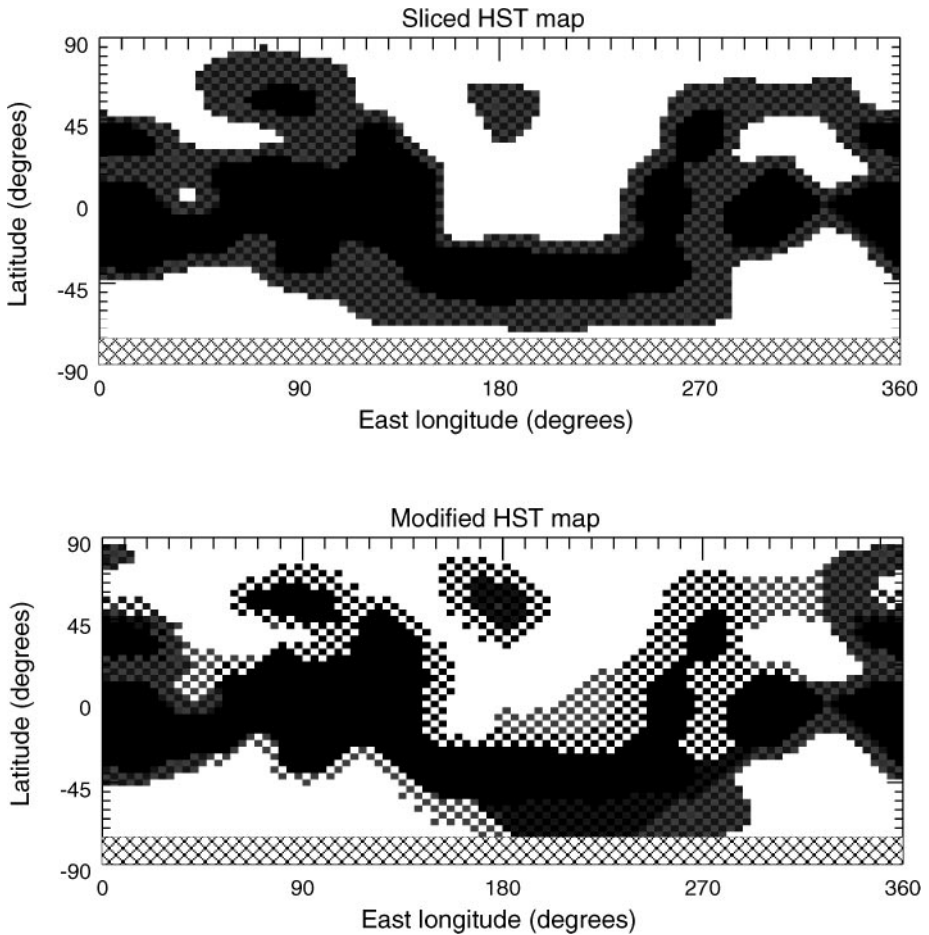


Figure 9 Compositional models from Grundy & Buie (2001). The models are created by splitting reanalyzed HST maps into three albedo regimes corresponding to three compositional units. The white regions are N_2 -rich units with diluted CH_4 and CO , the gray areas are units of almost pure CH_4 , and the dark areas are units with tholins mixed with water ice. The hatched regions near the south pole are areas unobserved by HST or the spectra. In the modified HST map the units have been slightly rearranged in a checkerboard pattern that still reproduces the HST albedos to produce a better fit to the spectral modulation.

have been coming into view are bright and varied. The HST map of spring equinox albedos can be viewed with this picture in mind; the southern cap is mostly gone with a few residual bright areas remaining, the high northern latitudes are bright, and the mid-latitudes are extremely variable. The difficulty is now to reconcile this picture with the Grundy & Buie terrain model, which suggests that the bright

areas are predominantly N_2 frost. The existence of any N_2 at the equator, which is receiving the maximum insolation that any region of the surface ever receives, is a puzzle unless Pluto has such a large reservoir of N_2 that permanent equatorial frost bands persist, as seen in the high N_2 abundance described in Hansen & Paige (1996) simulations (Figure 5). But why, in this case, would the equator be variable rather than uniformly frosty? In a region of permanent ice, the segregation into extremely dark and extremely bright patches should perhaps be expected. Any spot that is darker than average will absorb more sunlight, heat up, sublime more N_2 , and continue to darken; spots brighter than average will be colder and provide a place for condensation of the neighboring N_2 , leading to an even brighter surface. At the equator, this process could be self-sustaining. The equatorial regions never achieve the permanent night conditions that would allow the dark regions to cool down and become frosty, nor the permanent sunlight conditions that would allow the bright regions to completely sublime away and become dark and warm. This picture does not explain the origin of the bright and dark patches, but suggests that once albedo variations begin they can be exaggerated by volatile transport processes.

The final mystery remaining in this straw-man model, then, is the existence of patches of N_2 in the southern regions. Perhaps these are the last remnants of a once extensive south polar cap, and these particular spots have persisted longer because they began intrinsically bright (or other regions at their latitude were intrinsically dark). Perhaps these regions are telling us that the models for either seasonal change or surface composition on the surface of Pluto are flawed. To complete the model, the former is assumed, but this question remains open.

From this straw-man model, the near-future appearance of Pluto can be predicted. On a short-term timescale, the only clear predictions are that the southern N_2 patches should sublime quickly (or perhaps should even be gone by now), and the northern regions coming into continuous sunlight should be bright. Such changes should be visible to current-epoch HST imaging. The overall effect on Pluto's albedo of these two competing processes is difficult to foresee, but the light curve of Pluto could get more extreme as the variable bright southern patches disappear because these spots tend to be correlated with dark areas on the equator. Over the longer term (approaching summer solstice in 2029), the extreme variability at the equator should moderate as frosts begin to impinge on the outer regions of the dark spots, and the relatively bright northern hemisphere comes into full view. These two effects should cause Pluto to get brighter and have a muted light curve, much like it was when observed in 1955.

In the end, the model derived is (perhaps unsurprisingly) similar to the first-order expectations stated at the beginning, modified by some semipermanent albedo markings on Pluto's surface and some areas of permanent frost at the equator. While this model may prove to be overly simplistic or just plain wrong, the HST maps combined with rotational photometry and spectroscopy provide powerful but incomplete information on the surface composition and its variability on Pluto. A more complete understanding awaits better information on the compositional distribution of material on Pluto's surface, which will have to wait for either

detailed spectral mapping from a spacecraft flyby and/or infrared imaging and resolved spectroscopy from the next planned generation of 30+ meter ground-based telescopes. In the meantime, continued rotationally resolved spectra and photometry, along with a next epoch of HST imaging, are the best tools currently able to provide some insight into this difficult but fascinating problem.

CHARON, ITS SURFACE COMPOSITION, AND ITS RELATIONSHIP TO OTHER ICY BODIES

Charon orbits Pluto so closely that, even though it is only five times fainter than Pluto, Charon's existence was not discovered until 1978 (Christy 1978). Even after discovery, the small separation between the two objects made separate study of them difficult. The first knowledge of the composition of Charon came from observations of the mutual Pluto-Charon eclipses between 1985 and 1990. In a series of eclipses, Pluto and Charon were observed together, and then Pluto was observed separately as it completely occulted Charon. Subtraction of the two observations then yielded the brightness of Charon. By performing these observations at a small number of wavelengths over the near-infrared region, a low-resolution spectrum of Charon was synthesized. These spectra showed evidence for a surface covered in water ice, much like the icy satellites of the giant planets, and unlike the surface of Pluto (Marcialis et al. 1987, Buie et al. 1987). The possible existence of materials other than water ice was also considered, but the crudeness of the data did not permit any resolution of the issue (Roush 1994, Cruikshank et al. 1997).

The composition of Charon is possibly an important state in the transition from objects like the large icy satellites of the outer planets to the freely orbiting KBOs. Charon has some similarities to the icy satellites. With a radius of 593 km, it is approximately the same size as the mid-sized Saturnian satellites Tethys (~530 km) and Dione (560 km) and the Uranian satellites Ariel (~580 km) and Umbriel (585 km). Charon's visible albedo of 0.40 (Buie et al. 1990) is much lower than those of Tethys and Dione (0.9 and 0.7) and higher than the dark Umbriel (0.19), but closely matched to Ariel (0.35). The largest currently known KBO, (20000) Varuna, is only slightly smaller, with a radius of 450_{-75}^{+65} km, but is significantly darker, with a measured albedo of only $7.0_{-0.17}^{+3.0}$ % (Jewitt et al. 2001).

The proto-Charon presumably formed in a very different environment from the icy satellites. The regions of the solar nebula that were collapsing to form the giant planets were significantly higher in pressure than the surrounding nebula, leading to enrichments of CH₄ relative to CO and to NH₃ relative to N₂ (Prinn 1993). Thus, the first expectation should be that Pluto and Charon should be more abundant in CO and N₂ than CH₄ and NH₃, which should be more abundant on the icy satellites. The high N₂ abundance of Pluto is in accord with this scenario, but the existence of CH₄ on the body shows that this simple picture cannot account for all of the compositional variation. Nonetheless, significant compositional differences are expected between objects formed in high-pressure proto-giant-planetary environments and

objects formed in the lower-pressure solar nebula, and by comparing Charon to the similarly sized icy satellites, insight might be gained into these differing regions of the solar nebula.

Charon Spectroscopy

Recent improvements to ground- and space-based telescopes now allow direct spectroscopy of the individual members of the Pluto-Charon system. Spectra of the surface of Charon have now been obtained from the Keck telescope, Subaru telescope, and HST (under two independent programs) (Figure 10). The spectra all differ in signal-to-noise, resolution, reliability, and Charon longitude. A careful understanding of these factors allows us to use the combined information to make a best estimate of the surface composition of Charon. The first published results (Brown & Calvin 2000), which were obtained at the Keck telescope at a Charon longitude of 90° (Charon's longitude is defined such that the sub-Pluto longitude is at 0° and longitudes increase clockwise looking down on the rotational pole), are discussed first, and then these are compared with results from the other telescopes.

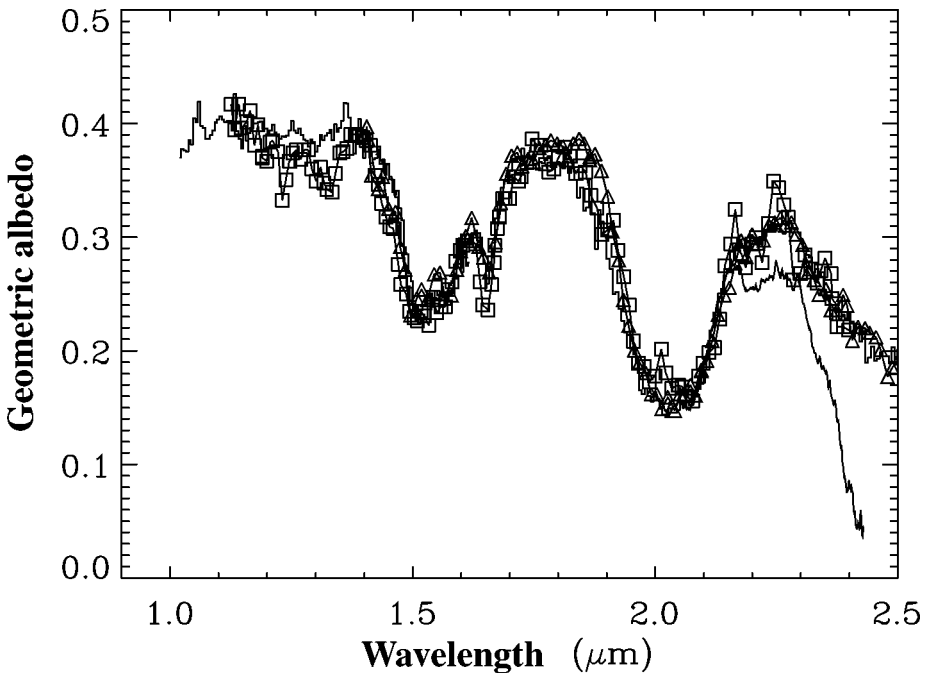


Figure 10 Four independent spectra of Charon. The Keck spectrum of Brown & Calvin (2000) is shown as the histogram, the Subaru spectrum of Nakamura et al. (2000) is shown as a solid line, the HST spectra of Buie & Grundy (2000) and Dumas et al. (2001) are shown as squares and triangles, respectively. The two ground-based spectra have been scaled to match the absolute albedo levels of the HST spectra.

CRYSTALLINE WATER ICE As previously found from the mutual event observations, the spectrum of Charon is dominated by the 1.5- and 2.0- μm absorptions of water ice. With the previous crude spectra from the mutual events, the form of the water ice could not be discerned. The new spectra clearly show the additional small 1.65- μm absorption feature redward of the main 1.5- μm absorption that occurs in crystalline water ice but not in amorphous water ice (Fink & Sill 1982). The existence of ice in crystalline form at Charon is unexpected. At temperatures below 70 K, amorphous ice is the expected stable form for water ice (Jenniskens et al. 1998), and crystalline water ice is turned into amorphous form under bombardment from solar ultraviolet radiation on short timescales (Kouchi & Kuroda 1990).

Many of the same arguments for the expectation of amorphous ice on Charon hold for the satellites of Uranus, yet spectroscopy shows that the water ice on these satellites is also in crystalline form (Grundy et al. 1999). The mechanism causing the ice to be in crystalline form on Charon is likely to be the same on these other satellites.

One method of making water ice condense into crystalline form rather than amorphous even at low temperatures is to condense it slowly. At Charon's temperature, condensation at a rate of approximately 1 cm h^{-1} or slower will preserve the ice in crystalline form (Jenniskens et al. 1998). Brown & Calvin (2000) suggest that crystalline ice on the surface of Charon could be caused by continuous micrometeorite impact vaporization and the subsequent recondensation of crystalline ice. If this process occurs faster than the radiation-induced transformation to amorphous ice, the surface could be maintained in crystalline form. Without knowing the timescale for either micrometeorite bombardment or irradiation amorphization, it is difficult to determine the feasibility of this scenario. Nonetheless, the presence of crystalline water ice on all of the well-studied icy satellites in the outer Solar System confirms that some ubiquitous mechanism such as impacts is likely responsible.

ABSORPTION AT 2.2 μm Models consisting of only crystalline water ice and a dark spectrally neutral material reproduce all of the major features of the Keck spectrum of Charon except at the 2.2- μm peak and longward of 2.3 μm . Figure 11 shows a spectral model consisting of spatially distinct regions of water ice and dark material using the optical constants of water ice at a temperature of 50 K measured by Grundy & Schmitt (1998) and assuming the reflectance spectroscopic theory of Hapke (1993). The model shown has a 70% coverage of water ice, with grain sizes of 30 μm and a 30% coverage by a dark spectrally neutral material, but because there are few water bands to fit and their absorption strengths do not vary considerably, many different combinations of grain size and water ice coverage would fit the data equally well. No grain size or combination of grain sizes can make the model fit the 2.2- μm region, however.

Before considering possible causes of this absorption, all spectra are compared to determine the reality of the feature. Buie & Grundy (2000) obtained spectra at Charon longitudes of 43, 131, 230, and 318° with HST. The spectral resolution is approximately 50% lower than that of the Keck observations,

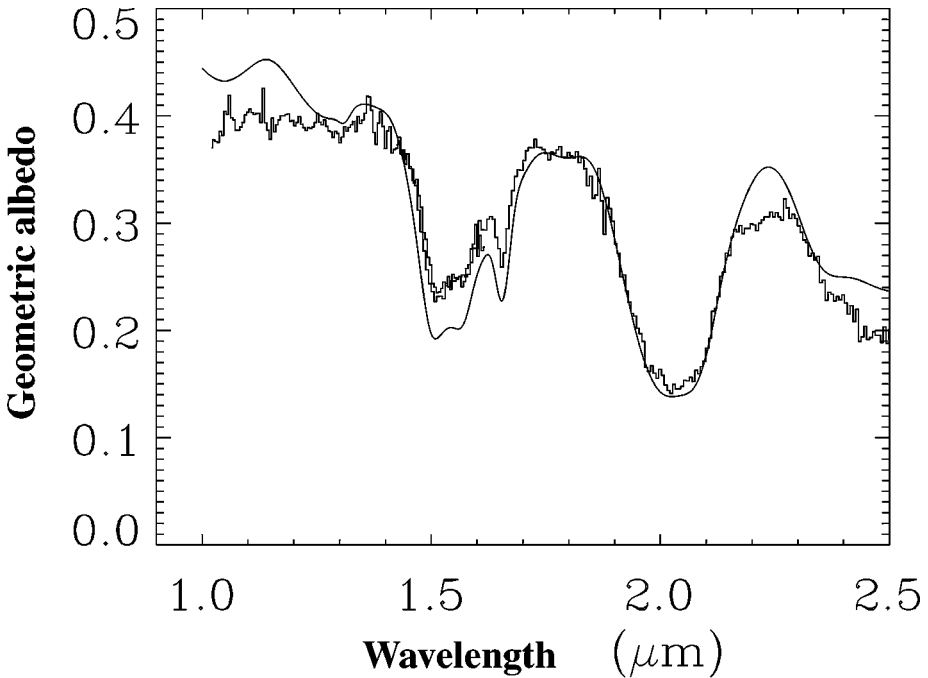


Figure 11 The Keck spectrum of Charon compared with the best-fit water ice model. Although different grain sizes and ice abundances can lead to adequate fits of most regions of the spectrum, the mismatch at $2.2 \mu\text{m}$ and beyond $2.3 \mu\text{m}$ cannot be explained by water ice.

whereas the signal-to-noise ratio per pixel is generally similar to that of the Keck observations (in regions where terrestrial water vapor is highly absorbing, however, the signal-to-noise ratio of the Keck spectra is lower). Figure 12 compares the Keck and HST spectra and shows that the two spectra are generally consistent but that the lower resolution of the HST spectrum partially masks the sharp appearance of the $2.2\text{-}\mu\text{m}$ absorption and instead makes it appear that there is a broad or continuous absorption beyond $\sim 2.0 \mu\text{m}$. The more recent results from the Dumas et al. (2001) HST observations and the Nakamura et al. (2000) Subaru observations, however, confirm the presence of a distinct $2.2\text{-}\mu\text{m}$ absorption rather than a broad absorption. Dumas et al. suggest that the leading hemisphere of Charon (longitude of 90°) has a distinct $2.2\text{-}\mu\text{m}$ absorption, whereas the trailing hemisphere (longitude of 270°) does not. The signal-to-noise ratio of their spectrum is lower than that of the Keck or other HST spectrum (the instrument and therefore resolution is identical to Buie & Grundy), but the absorption appears at the level of several standard deviations. One explanation for the discrepancy between the two HST spectra is that the $2.2\text{-}\mu\text{m}$ absorption is strongest on the leading hemisphere of Charon. The Buie

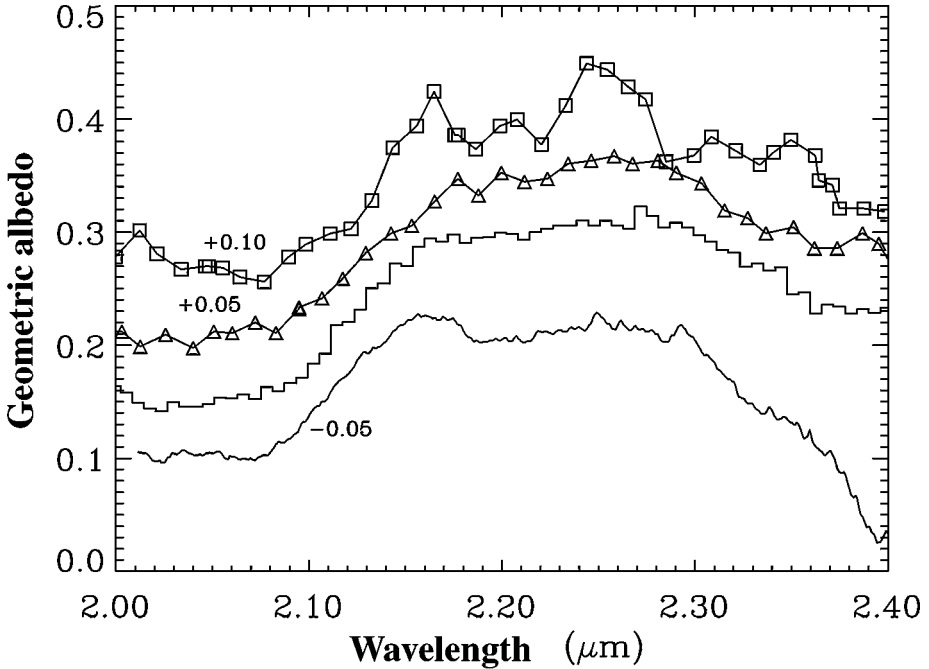


Figure 12 A comparison of the 2.2- μm region of the four Charon spectra. The spectra are distinguished as in previous figures and are offset by the labeled amounts for clarity. The spectral resolution of the observation increases from bottom to top (with the top two HST spectra identical in resolution). The two higher-resolution ground-based spectra show evidence of the 2.2- μm absorption, as does the top HST spectrum, though the signal-to-noise ratio of that spectrum is low.

& Grundy HST observations were averaged over four longitudes (none of which was centered on 90°) so a 2.2- μm absorption occurring only at 90° would have been degraded. The one Buie & Grundy spectrum taken at 230° shows a hint of the absorption, though the signal-to-noise ratio of the individual spectrum is not high. The Subaru spectrum is of the same 90° longitude as the Keck and Dumas et al. HST spectra, and at much higher spectral resolution, clearly shows a distinct 2.2- μm absorption, though the overall calibration of the Subaru spectrum does not agree with any of the other three spectra. The distinct 2.2- μm absorption, first reported by Brown & Calvin, thus appears to be a real feature of Charon's spectrum, and is possibly strongest on the leading hemisphere.

Brown & Calvin, Buie & Grundy, and Dumas et al. each considered different possibilities for absorption in this spectral region. Buie & Grundy sought a substance with a broad absorption beyond $2.0 \mu\text{m}$, but because the higher resolution spectra show that the absorption is distinct rather than broad, those substances are not considered here. Brown & Calvin considered all ices previously observed

on or proposed for Solar System bodies and the interstellar medium. In addition, they considered all other simple combinations of H, C, N, and O, the four most abundant molecule-forming elements in the Solar System, and more complex hydrocarbons such as tholins and kerogens, which have long been considered plausible causes for the low albedo surfaces in the outer Solar System. Given the recent reports of hydrated minerals on the icy satellites of Jupiter (McCord et al. 1999) they also considered all common terrestrial rock-forming minerals. Of all of these possible constituents, only HCN, NH_3 , $\text{NH}_3 \cdot 2\text{H}_2\text{O}$ (ammonia hydrate), and aluminosilicate clays have distinct absorption features in the $2.2\text{-}\mu\text{m}$ region. Dumas et al. came to similar conclusions about the nature and shape of the absorption and considered in detail HCN and ammonia hydrate. Figure 13 shows the spectra of these materials and compares them to the Keck and higher-resolution Subaru spectra of Charon. Detailed modeling by Brown & Calvin showed that the clays have a band center shifted too far blueward to fit Charon's spectrum and that the HCN absorption is far too broad to account for the distinct $2.2\text{-}\mu\text{m}$ absorption.

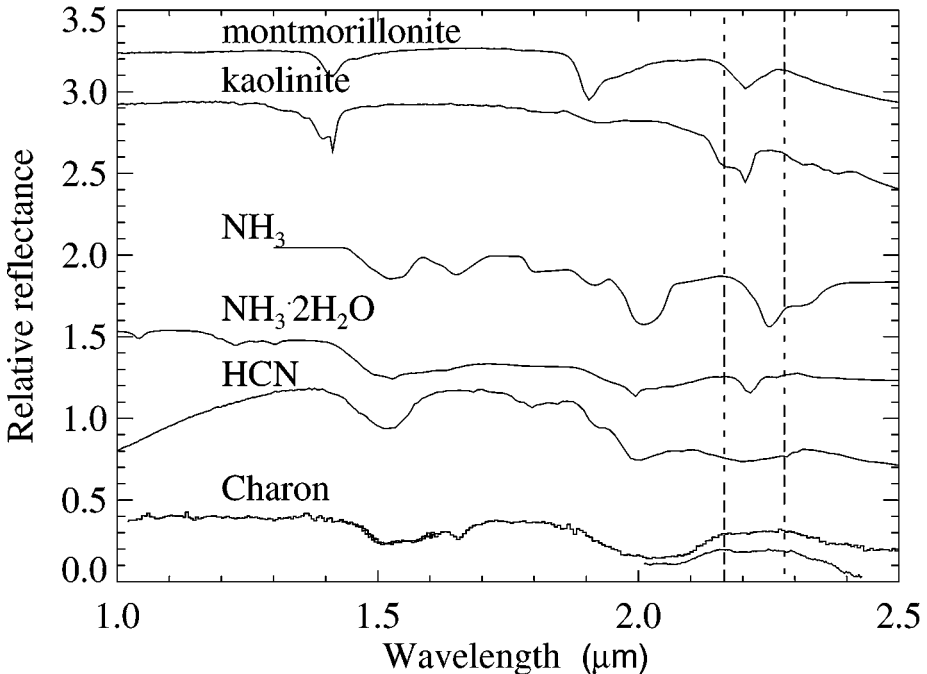


Figure 13 Spectra of clays and ices with absorptions at $2.2\text{ }\mu\text{m}$ (from Brown & Calvin 2000), compared with the Keck and Subaru Charon spectra. Each spectrum (except for Subaru) is offset by 0.5 in reflectance from the lower spectrum. The dashed vertical lines show the extent of the observed absorption. None of the individual spectra matches the absorption, but a combination of ammonia and ammonia hydrate provides a good fit to the spectra.

Dumas et al. reach a similar conclusion regarding the width of the HCN absorption (though at their lower spectral resolution the differences are harder to distinguish, and indeed, even the aluminosilicate clays, which they did not consider, would provide an adequate fit). The higher-resolution spectrum from Subaru agrees well with the moderate-resolution spectrum from Keck (other than uncertain intensity calibration) and conclusively shows that the absorption produced by the clays is insufficient to explain the 2.2- μm absorption and that the width and shape of HCN cannot explain the absorption.

The only substance found by Brown & Calvin, Buie & Grundy, or Dumas et al. that can adequately match the width and shape of the 2.2- μm absorption is a combination of pure ammonia and ammonia hydrate. Ammonia hydrate by itself has a 2.2- μm absorption similar to those of the clays, and thus cannot explain the spectrum of Charon, but pure ammonia ice has an absorption shifted redward from the hydrated form. A combination of the two ices creates an absorption with the correct width and shape to explain the feature observed on Charon. In addition, all other absorption features of ammonia and ammonia hydrate correspond to strong water ice absorptions, so the spectral signature of water ice is not disturbed elsewhere, as seen in the preferred Brown & Calvin model consisting of water ice (44% surface coverage), ammonia (1%), ammonia hydrate (24%), and a dark neutral material (32%) (Figure 14). Encouragingly, this spectral model, which was developed to match the moderate resolution Keck spectrum, does an excellent job of reproducing the detailed structure seen in the higher-resolution Subaru spectrum. The Dumas et al. HST spectrum is better fit by a model that does not include the extra 2.25- μm absorption from pure ammonia ice, but the agreement between the Keck and Subaru data on the presence of 2.25- μm absorption suggests that the discrepancy between those spectra and that of Dumas et al. is caused by the lower signal-to-noise ratio of the Dumas et al. data.

Although the match of the spectra with ammonia ices is good, the identification is not unique. It remains possible that some not yet considered substance is the true cause of the absorption. High priority should be given to obtaining high-quality high-resolution spectra of the full near-infrared range for both hemispheres of Charon. The ammonia ices have subtle features throughout this spectral range that are difficult to identify in a single spectrum, but might become obvious by their changes from one hemisphere to the other. In the meantime, ammonia ices are a good tentative match. Their consequences are considered below.

AMMONIA ICES Ammonia should not be particularly abundant in the low-pressure solar nebula; equilibrium chemistry drives nitrogen toward N_2 . Some ammonia must remain, however, because comet Hale-Bopp was seen to have an ammonia abundance of $\sim 1\%$ compared with water (Bird et al. 1997). Assuming the proto-Pluto and Charon were formed in similar regions of the nebula as comets, their ammonia abundances should be similarly small. Yet, the best-match Charon model has a surface approximately one quarter covered with ammonia ices. This discrepancy suggests that ammonia can be at best a surficial covering on the satellite.

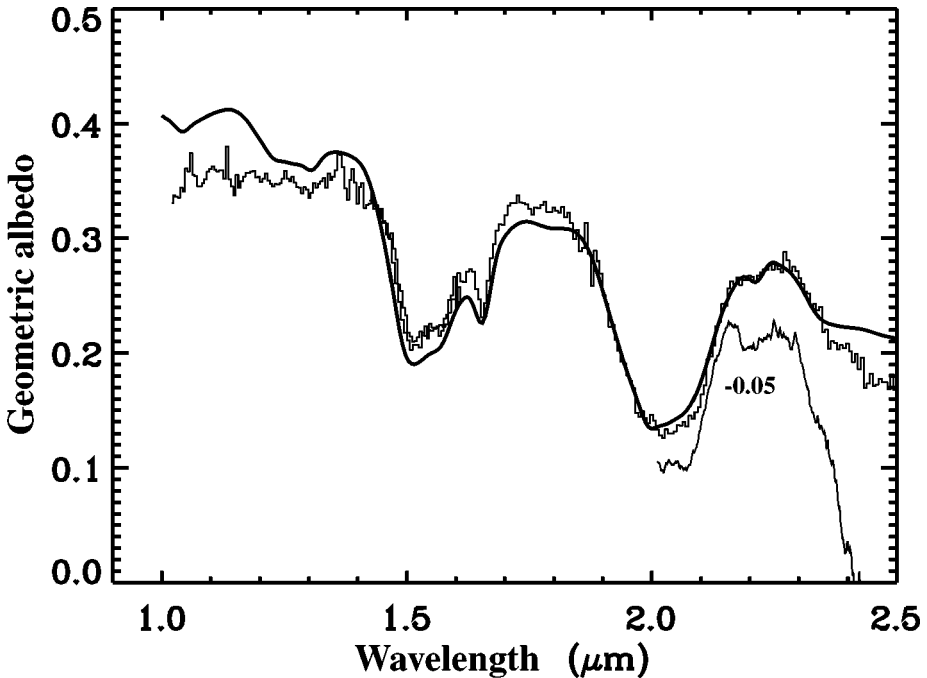


Figure 14 The Keck (*histogram*) and Subaru (*offset line*) spectra of Charon compared with an ammonia ice model. The model is a combination of water ice (44% surface coverage), ammonia (1%), ammonia hydrate (24%), and a dark neutral material (32%). The model provides a good match to the 2.2- μm region; no attempt was made to fit the absolute albedo levels below 2.0 μm . Note that the absorption past 2.35 μm , seen in the spectra of most icy satellites, is still not fit.

A surficial covering can be of either exogenic or endogenic origin. Based on the potential leading-trailing asymmetry of the appearance of ammonia and the analogy with similar asymmetries seen in the Galilean and Saturnian satellites, Dumas et al. explore an exogenic possibility related to nitrogen on Pluto. They suggest that the escape and subsequent ionization of N_2 from Pluto could lead to ions becoming implanted preferentially in the leading hemisphere of Charon. This ion bombardment could also provide the energy required to form ammonia from the water ice and nitrogen ion surface. That such chemistry is possible has not been demonstrated; Lanzerotti et al. (1984) have shown that energetic ion bombardment leads to the breakdown, rather than build up, of ammonia with water ices. In addition, there is no reason to expect ions to bombard the surface of Charon at high energy or preferentially on one side. At Jupiter, for example, a primary energy source for the ions is the fact that they are created far out in Jupiter's magnetic field (dominantly at the location of Io) and are quickly accelerated to

corotate with the much more quickly moving magnetic field. These high-energy ions then preferentially bombard the trailing hemispheres of the satellites. On Pluto, in contrast, the ion source would be the planet itself, so little acceleration would need to happen. In addition, Charon orbits Pluto at the same speed as that of Pluto, and thus any magnetic field, rotates. Any ions trapped in the magnetosphere would therefore be going at approximately the same speed as Charon and would neither preferentially bombard one hemisphere nor implant with high energy. Although this mechanism has some appeal because of the large potential source of nitrogen from Pluto, it appears unlikely.

Brown & Calvin suggest an endogenic origin for ammonia. They note that ammonia ices have long been predicted to be an important component of icy satellites in the outer Solar System (Lewis 1971). The discovery by Voyager that even relatively small satellites, with interiors too cold to melt pure ice, have had complex geologic histories led to the realization that ammonia-water mixtures, because of their lower melting temperatures and higher viscosities than pure water, may be responsible for the activity seen on these bodies (Stevenson 1982, Schenk 1991, Kargel et al. 1991, Durham et al. 1993). Both Ariel (Figure 15) and Umbriel, the two satellites most similar in size and albedo to Charon, exhibit a variety of surface geologic units including cratered plains and fault valleys whose floors suggest infilling by erupted and flowing materials similar to terrestrial volcanic flows (Croft & Soderblom 1997). Charon should have experienced similar amounts of accretional heating in the past, in addition to any extra heating from tidal evolution of the Pluto-Charon system, so similar geological activity on this body is expected. The detection of ammonia ices on Charon would suggest that such ices play an important role in geological activity on icy bodies in the outer Solar System.

The major difficulties faced by the volcanic scenario are (a) that ammonia ices are destroyed by irradiation (Lanzerotti et al. 1984), so are not expected to exist on a satellite surface for an extended period of time; and (b) ammonia ices have never been detected on any of the other satellites for which they have been hypothesized to have caused geological activity. The apparent radiation destruction of ammonia is an important constraint; if ammonia is not stable in space environments, its apparent spectroscopic identification must be in error. Nonetheless, the same statement can be made about crystalline water, which should not be stable on the surface of Charon. Clearly, the understanding of the stability of ices in the space environment is incomplete, so this constraint should be taken with caution. For example, given the long times available on planetary surfaces, the products of ammonia irradiation possibly eventually recombine to reform the original ammonia.

The second difficulty is that ammonia has not been detected on any of the other icy satellites on which it has been hypothesized to have been an important geological driver. Brown & Calvin show, however, that ammonia ices are sufficiently volatile that they would have long ago evaporated from the surface of any of the icy satellites except for those of Uranus. The highest-quality published spectra of the Uranian satellites (Grundy et al. 1999) have sufficiently low signal-to-noise in the 2.2- μm region that the presence of a 2.2- μm absorption cannot be determined



Figure 15 The surface of Ariel, a satellite of Uranus with size, spectrum, and albedo similar to Charon. The smooth areas inside the large crack systems are hypothesized to be due to ammonia-water cryovolcanic resurfacing. The same processes could have operated on Charon, possibly leading to a similar-looking surface.

(see below). An additional observational priority should be the acquisition of high-resolution, high signal-to-noise spectra of the Uranian satellites to allow detailed comparison with the spectrum of Charon.

Relationship of Charon to Other Icy Bodies

This section began with the philosophy that because of its formation in the low-pressure solar nebula, Charon should be more like the dark KBOs in composition than the icy satellites. Spectroscopically, however, the opposite appears to be true.

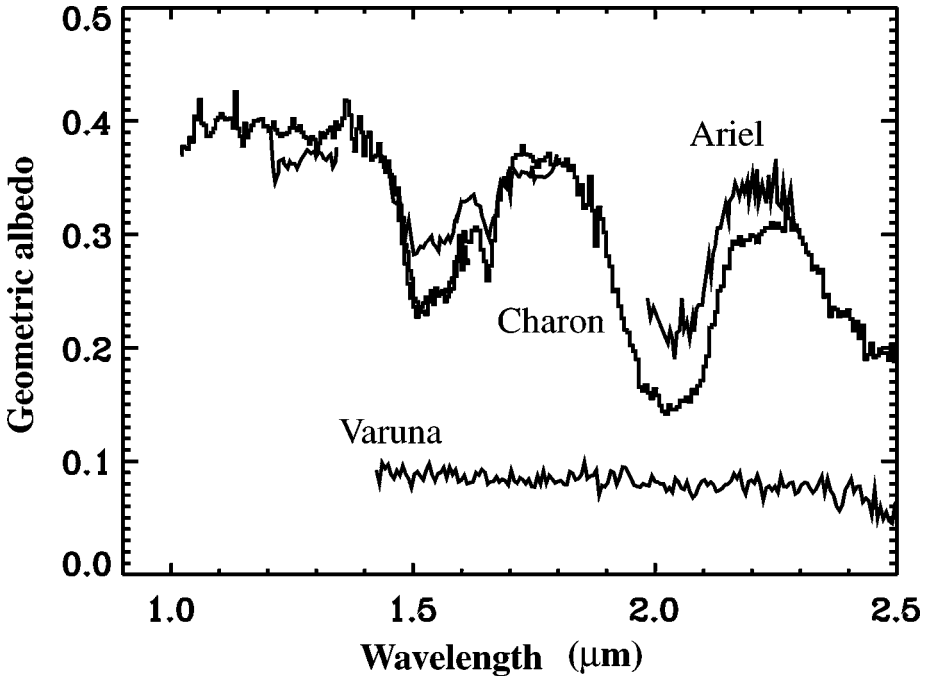


Figure 16 A comparison of the spectra of three like-sized bodies: Charon (Brown & Calvin 2000), Ariel (Grundy et al. 1999), and the large KBO Varuna (unpublished data of M. E. Brown). The two icy satellites are similar in absolute albedo level and in water ice absorption level. The KBO is significantly darker, with an essentially featureless spectrum.

Figure 16 compares the reflection spectra of Charon, Ariel, and the KBO (20000) Varuna, all of which are approximately the same size. The spectrum of Ariel is a close match in both features and albedo level to that of Charon. The spectrum of Varuna is featureless and significantly darker. For many years speculation has suggested that as KBOs approach the size of Charon they should begin to approach the spectrum and albedo of Charon also, but the features of Varuna clearly dash these expectations.

KBOs are thought to be dark because of the darkening of organic ices exposed to irradiation (similar to that discussed above for tholins; Strazzulla & Palumbo 1998). But the icy satellites have been exposed to an irradiation environment similar to that of the Kuiper belt, so why are they bright? Presumably, the answer lies in the different chemistry in the satellite-forming nebula discussed above. The elevated temperatures and pressures must have removed whatever organic components are responsible for darkening, allowing the icy satellites to stay bright even under irradiation. Why, then, is Charon, which presumably formed in regions similar to those in which KBOs formed, also bright? The only reasonable hypothesis seems to be that the heat generated in the Charon-forming collision was sufficient to remove

those same darkening agents on Charon. This hypothesis predicts that other binary KBOs formed in catastrophic collisions should have similarly elevated albedos.

These conclusions come mainly from the albedo measurements of one large KBO. Clearly, measurement of a significant sample of KBO albedos should be a high priority for understanding the chemistry of the outer solar system and the relationship of Charon and the icy satellites to the KBOs.

OUTSTANDING QUESTIONS

Many questions remain relating to the topics covered above:

- What processes, other than resonance capture, operated in the early outer Solar System and gave Pluto and the KBOs their large eccentricities and inclinations?
- What mechanism was mainly responsible for clearing out the Solar System; erosion or dynamics (or something else)? How many Pluto-sized bodies existed in the early Solar System? Do any others still exist?
- Can simulations that include the three main compositional units reproduce the observations of seasonal variation on Pluto?
- How do the bright and dark spots relate to spectroscopic compositional units? How much does the surface change? Do long-lived N₂ deposits exist near the equator? Do N₂ deposits still survive in the southern region?
- What is the surface composition of Charon? How is it spatially distributed?
- Why is outer Solar System surface water ice in crystalline form?
- Do ammonia ices exist on other satellites? Are they stable against irradiation?
- Why is Charon bright and like-sized KBOs dark? What is the distribution of KBO albedos?

Answers to these questions are within reach of the current generation of Pluto explorers and will provide insights and answers to questions of the formation, seasonal variability, and composition of Pluto and Charon.

Visit the Annual Reviews home page at www.annualreviews.org

LITERATURE CITED

- Bird MK, Huchtmeier WK, Gensheimer P, Wilson TL, Janardhan P, Lemme C. 1997. Radio detection of ammonia in comet Hale-Bopp. *Astron. Astrophys.* 325:L5–8
- Brown ME. 2001. The inclination distribution of the Kuiper belt. *Astron. J.* 121:2804–14
- Brown ME, Calvin WM. 2000. Evidence for crystalline water and ammonia ices on Pluto's satellite Charon. *Science* 287:107–9
- Brown RH, Cruikshank DP. 1997. Determination of the composition and state of icy surfaces in the outer solar system. *Annu. Rev. Earth Planet. Sci.* 25:243–77
- Buie MW, Cruikshank DP, Lebofsky LA,

- Tedesco EF. 1987. Water frost on Charon. *Nature* 329:522–23
- Buie MW, Fink U. 1987. Methane absorption variations in the spectrum of Pluto. *Icarus* 70:483–98
- Buie MW, Grundy WM. 2000. The distribution of physical state of H₂O on Charon. *Icarus* 148:324–39
- Buie MW, Tholen DJ, Horne K. 1992. Albedo maps of Pluto and Charon—initial mutual event results. *Icarus* 98:211–27
- Buie MW, Young EF, Binzel RP. 1997. Surface appearance of Pluto and Charon. See Stern & Tholen 1997, pp. 269–94
- Christy JW. 1978. *IAU Circ.* No. 3241
- Cohen CJ, Hubbard EC. 1965. Librations of the close approaches of Pluto to Neptune. *Astron. J.* 70:10–13
- Croft SK, Soderblom LA. 1991. Geology of Uranian satellites. In *Uranus*, ed. J Bergstralh, E Miner, M Matthews. Tucson: Univ. Ariz. Press. 561 pp.
- Cruikshank DP, Pilcher CD, Morrison D. 1976. Pluto: evidence for methane frost. *Science* 194:835–37
- Cruikshank DP, Roush TL, Bartholomew NJ, Geballe TR, Pendleton YJ, et al. 1998. The composition of Centaur 5146 Pholus. *Icarus* 135:389–407
- Cruikshank DP, Roush TL, Moore JM, Sykes M, Owen TC, et al. 1997. The surface of Pluto and Charon. See Stern & Tholen 1997, pp. 221–68
- Cruikshank DP, Roush TL, Owen TC, Geballe TR, de Bergh C, et al. 1993. Ices on the surface of Triton. *Science* 261:742–45
- Douté S, Schmitt B, Quirico E, Owen TC, Cruikshank DP, et al. 1999. Evidence for methane segregation at the surface of Pluto. *Icarus* 142:421–44
- Drish WF, Harmon R, Marcialis RL, Wild W. 1995. Images of Pluto generated by matrix lightcurve inversion. *Icarus* 113:360–86
- Dumas C, Terrile RJ, Brown RH, Schneider G, Smith BA. 2001. Hubble Space Telescope NICMOS spectroscopy of Charon's leading and trailing hemispheres. *Astron. J.* 121:1163–70
- Duncan NJ, Levison HF, Budd SM. 1995. The dynamical structure of the Kuiper belt. *Astron. J.* 110:3073–94
- Durham WB, Kirby SH, Stern LA. 1993. Flow of ices in the ammonia-water system. *J. Geophys. Res.* 98:17667–82
- Elliot JL, Dunham EW, Bosh AS, Slivan SM, Young LA, et al. 1989. Pluto's atmosphere. *Icarus* 77:148–70
- Fink U, Sill GT. 1982. The infrared spectral properties of frozen volatiles. In *Comets*, ed. L Wilkening. Tucson: Univ. Ariz. Press. 172 pp.
- Gomes RS. 2000. Planetary migration and plutino orbital inclinations. *Astron. J.* 120:2695–707
- Grundy WM, Buie MW. 2001. Distribution and evolution of CH₄, N₂, and CO ices on Pluto's surface: 1995 to 1998. *Icarus*. 2: 248–63
- Grundy WM, Buie MW, Stansberry JA, Spencer JR, Schmitt G. 1999. Near-infrared spectra of icy outer Solar System surfaces: remote determination of H₂O ice temperatures. *Icarus* 132:536–49
- Grundy WM, Fink U. 1996. Synoptic CCD spectrophotometry of Pluto over the past 15 years. *Icarus* 124:329–43
- Grundy WM, Schmitt B. 1998. The temperature-dependent near-infrared absorption spectrum of hexagonal H₂O ice. *J. Geophys. Res.* 103:25809–22
- Hahn JM, Malhotra R. 1999. Orbital evolution of planets embedded in a planetesimal disk. *Astron. J.* 117:3041–53
- Hansen C, Paige DA. 1996. Seasonal nitrogen cycles on Pluto. *Icarus* 120:247–65
- Hapke B. 1993. *Theory of Reflectance and Emittance Spectroscopy*. New York: Cambridge Univ. Press
- Hubbard WB, Hunten DM, Deiters SW, Hill KM, Watson RD. 1988. Occultation evidence for an atmosphere on Pluto. *Nature* 336:452–54
- Jenniskens P, Blake DF, Kouchi A. 1998. Amorphous water ice. A solar system material. In *Solar System Ices*, ed. B Schmitt, C de Bergh, M Festou. Dordrecht: Kluwer. 139 pp.

- Jewitt D. 1999. Kuiper belt objects. *Annu. Rev. Earth Planet. Sci.* 27:287–312
- Jewitt D, Aussen H, Evans A. 2001. The size and albedo of the Kuiper-belt object (20000) Varuna. *Nature* 411:446–47
- Kargel JS, Croft SK, Lunine JI, Lewis JS. 1991. Rheological properties of ammonia-water liquids and crystal-liquid slurries—Planetological applications. *Icarus* 89:93–112
- Kenyon SJ, Luu JX. 1998. Accretion in the early Kuiper Belt. I. Coagulation and velocity evolution. *Astron. J.* 115:2136–60
- Kenyon SJ, Luu JX. 1999. Accretion in the early Kuiper Belt. II. Fragmentation. *Astrophys. J.* 118:1101–19
- Khare BN, Sagan C, Arakawa ET, Suits F, Callcott TA, Williams MW. 1984. Optical constants of organic tholins produced in simulated Titanian atmosphere: from soft X-ray to microwave frequencies. *Icarus* 60:127–37
- Kouchi A, Kuroda T. 1990. Amorphization of cubic ice produced by ultraviolet-irradiation. *Nature* 344:134–35
- Lanzerotti LJ, Brown WL, Marcantonio KJ, Johnson RE. 1984. Production of ammonia-depleted surface-layers on the Saturnian satellites by ion sputtering. *Nature* 312:139–40
- Lellouch E, Laureijs R, Schmitt B, Quirico E, de Berg C, et al. 2000. Pluto's non-isothermal surface. *Icarus* 147:220–50
- Lewis JS. 1971. Satellites of the outer planets: their physical and chemical nature. *Icarus* 15:174–80
- Lissauer JJ, Pollack JB, Wetherill GW, Stevenson DJ. 1995. Formation of the Neptune system. In *Neptune and Triton*, ed. D Cruikshank. Tucson: Univ. Ariz. Press. 37 pp.
- Lissauer JJ, Stewart GR. 1993. Growth of planets from planetesimals. In *Protostars and Planets III*, ed. E Levy, J Lunine. Tucson: Univ. Ariz. Press. 1061 pp.
- Malhotra R. 1993. The origin of Pluto's peculiar orbit. *Nature* 365:819–21
- Malhotra R. 1995. The origin of Pluto's orbit—implications for the solar-system beyond Neptune. *Astron. J.* 110:420–29
- Malhotra R. 1998. Pluto's inclination excitation by resonance sweeping. *Lunar Planet. Sci. Conf.* 29:1476–77
- Malhotra R, Williams JG. 1997. Pluto's heliocentric orbit. See Stern & Tholen 1997, pp. 127–58
- Marcialis RL. 1988. A two-spot model for the surface of Pluto. *Astron. J.* 95:941–47
- Marcialis RL, Lebofsky LA. 1991. CVF spectrophotometry of Pluto—correlation of composition with albedo. *Icarus* 89:255–63
- Marcialis RL, Rieke GH, Lebofsky LA. 1987. The surface composition of Charon—tentative identification of water ice. *Science* 237:1349–51
- McCord TB, Hansen GB, Matson DL, Johnson TV, Crowley JK, et al. 1999. Hydrated salt minerals on Europa's surface from the Galileo near-infrared mapping spectrometer (NIMS) investigation. *J. Geophys. Res.* 104:11827–51
- Morbidelli A, Valsecchi GB. 1997. Neptune scattered planetesimals could have sculpted the primordial Edgeworth-Kuiper belt. *Icarus* 128:464–68
- Nakamura R, Sumikawa S, Ishiguro M, Mukai T, Iwamuro F, et al. 2000. Subaru infrared spectroscopy of the Pluto-Charon system. *Publ. Astron. Soc. Jpn.* 52:551–56
- Owen TC, Roush TL, Cruikshank DP, Elliot JL, Young LA, et al. 1993. Surface ices and the atmospheric composition of Pluto. *Science* 261:745–48
- Petit JM, Morbidelli A, Valsecchi GB. 1999. Large scattered planetesimals and the excitation of the small body belts. *Icarus* 141:367–87
- Prinn RG. 1993. Chemistry and evolution of gaseous circumstellar disks. In *Protostars and Planets III*, ed. E Levy, J Lunine. Tucson: Univ. Ariz. Press. 1005 pp.
- Roush TL. 1994. Charon—more than water ice. *Icarus* 108:243–54
- Quirico E, Schmitt B. 1997. A spectroscopic study of CO diluted in N⁻² ice: applications for Triton and Pluto. *Icarus* 128:181–88
- Quirico E, Schmitt B, Bini R, Salvi PR. 1996. Spectroscopy of some ices of astrophysical

- interest: SO_2 , N^{-2} , and $\text{N}^{-2}:\text{CH}_4$ mixtures. *Planet. Space Sci.* 44:973–86
- Schenk PM. 1991. Fluid volcanism on Miranda and Ariel—flow morphology and composition. *J. Geophys. Res.* 96:1887–906
- Spencer JR, Buie MW, Bjoraker GL. 1990. Solid methane on Triton and Pluto—3 μm to 4 μm spectrophotometry. *Icarus* 88:491–96
- Spencer JR, Stansberry JA, Trafton LM, Young EF, Binzel RP, Croft SK. 1997. Volatile transport, seasonal cycles, and atmospheric dynamics on Pluto. See Stern & Tholen 1997, pp. 435–74
- Stern SA. 1991. On the number of planets in the outer solar-system—evidence of a substantial population of 1000 km bodies. *Icarus* 90:271–81
- Stern SA. 1992. The Pluto-Charon system. *Ann. Rec. Astron. Astrophys.* 30:185–233
- Stern SA. 1995. Collisional time scales in the Kuiper disk and their implications. *Astron. J.* 110:856–68
- Stern SA. 1996. On the collisional environment, accretion time scales, and architecture of the massive, primordial Kuiper belt. *Astron. J.* 112:1203–11
- Stern SA, Buie MW, Trafton LM. 1997a. HST high-resolution images and maps of Pluto. *Astron. J.* 113:827–35
- Stern SA, Colwell JE. 1997a. Accretion in the Edgeworth-Kuiper Belt: forming 100–1000 km radius bodies at 30 AU and beyond. *Astron. J.* 114:841–62
- Stern SA, Colwell JE. 1997b. Collisional erosion in the primordial Edgeworth-Kuiper Belt and the generation of the 30–50 AU Kuiper gap. *Astrophys. J.* 490:879–93
- Stern SA, McKinnon WB, Lunine JJ. 1997b. On the origin of Pluto, Charon, and the Pluto-Charon binary. See Stern & Tholen 1997, pp. 605–64
- Stern SA, Tholen DJ, eds. 1997. *Pluto and Charon*. Tucson: Univ. Ariz. Press. 729 pp.
- Stevenson DJ. 1982. Volcanism and igneous processes in small icy satellites. *Nature* 298:142–44
- Strazzulla G, Palumbo ME. 1998. Evolution of icy surfaces: an experimental approach. *Planet. Space Sci.* 46:1339–48
- Trafton LM, Hunten DM, Zahnle KJ, McNutt RL. 1997. Escape processes at Pluto and Charon. See Stern & Tholen 1997, pp. 475–522
- Trafton LA, Stern SA. 1996. Rotationally resolved spectral studies of Pluto from 2500 to 4800 angstrom obtained with HST. *Astron. J.* 112:1212–24
- Walker MF, Hardie RH. 1955. A photometric determination of the rotational period of Pluto. *Publ. Astron. Soc. Pac.* 67:224–31
- Young EF, Binzel RP, Crane K. 2001. A two-color map of Pluto's sub-Charon hemisphere. *Astron. J.* 121:552–61
- Young EF, Galdamez K, Buie MW, Binzel RP, Tholen DJ. 1999. Mapping the variegated surface of Pluto. *Astron. J.* 117:1063–76



CONTENTS

Frontispiece— <i>Vladimir Keilis-Borok</i>	xvi
EARTHQUAKE PREDICTION: STATE-OF-THE-ART AND EMERGING POSSIBILITIES, <i>Vladimir Keilis-Borok</i>	1
MODELING COMPLEX, NONLINEAR GEOLOGICAL PROCESSES, <i>Greg A. Valentine, Dongxiao Zhang, and Bruce A. Robinson</i>	35
DATING THE TIME OF ORIGIN OF MAJOR CLADES: MOLECULAR CLOCKS AND THE FOSSIL RECORD, <i>Andrew B. Smith and Kevin J. Peterson</i>	65
MODERN INTEGRATIONS OF SOLAR SYSTEM DYNAMICS, <i>A. Morbidelli</i>	89
IMPLICATIONS OF EXTRASOLAR PLANETS FOR UNDERSTANDING PLANET FORMATION, <i>Peter Bodenheimer and D.N.C. Lin</i>	113
SCALING OF SOIL MOISTURE: A HYDROLOGIC PERSPECTIVE, <i>Andrew W. Western, Rodger B. Grayson, and Günter Blöschl</i>	149
STREAMFLOW NECESSARY FOR ENVIRONMENTAL MAINTENANCE, <i>Peter J. Whiting</i>	181
PETROLOGY OF SUBDUCTED SLABS, <i>Stefano Poli and Max W. Schmidt</i>	207
GEODYNAMO SIMULATIONS—HOW REALISTIC ARE THEY?, <i>Gary A. Glatzmaier</i>	237
MODERN IMAGING USING SEISMIC REFLECTION DATA, <i>Michael C. Fehler and Lianjie Huang</i>	259
PRELUDE TO THE CAMBRIAN EXPLOSION, <i>James W. Valentine</i>	285
PLUTO AND CHARON: FORMATION, SEASONS, COMPOSITION, <i>Michael E. Brown</i>	307
GEOLOGIC STRUCTURE OF THE UPPERMOST OCEANIC CRUST CREATED AT FAST-TO INTERMEDIATE-RATE SPREADING CENTERS, <i>Jeffrey A. Karson</i>	347
VOLCANOES, FLUIDS, AND LIFE AT MID-OCEAN RIDGE SPREADING CENTERS, <i>Deborah S. Kelley, John A. Baross, and John R. Delaney</i>	385
MANTLE MIXING: THE GENERATION, PRESERVATION, AND DESTRUCTION OF CHEMICAL HETEROGENEITY, <i>Peter E. van Keken, Erik H. Hauri, and Chris J. Ballentine</i>	493

FOSSIL PLANTS AS INDICATORS OF THE PHANEROZOIC GLOBAL
CARBON CYCLE, *D.J. Beerling and D.L. Royer* 527

INDEXES

Subject Index 557
Cumulative Index of Contributing Authors, Volumes 20–30 583
Cumulative Index of Chapter Titles, Volumes 20–30 586

ERRATA

An online log of corrections to *Annual Review of Earth and Planetary Sciences* chapters (if any, 1997 to the present) may be found at <http://earth.annualreviews.org/>